

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 М. Никулин (M. Nikulin, Bordeaux, France)
профессор В. Ротарь (V. Rotar, San-Diego, USA)
профессор М. Финкельштейн (M. Finkelstein, Bloemfontein, South Africa)

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

академик РАН И. А. Соколов — главный редактор
профессор, д.ф.-м.н. С. Я. Шоргин — заместитель главного редактора
д.т.н. В. Н. Захаров д.ф.-м.н. В. И. Синицын
проф., д.ф.-м.н. А. И. Зейфман проф., д.т.н. И. Н. Синицын
проф., д.т.н. В. Д. Ильин проф., д.ф.-м.н. В. Г. Ушаков
проф., д.т.н. К. К. Колин к.ф.-м.н. А. К. Горшенин — отв. секретарь
проф., д.ф.-м.н. В. Ю. Королев к.ф.-м.н. С. А. Христочевский
проф., д.г.-м.н. Р. Б. Сейфуль-Мулюков

Редакция

профессор, д.г.-м.н. Р. Б. Сейфуль-Мулюков
к.ф.-м.н. Е. Н. Арутюнов
С. Н. Стригина

© Федеральный исследовательский центр «Информатика
и управление» Российской академии наук, 2018

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 28 № 4 Год 2018

СОДЕРЖАНИЕ

Автоматическое построение синтаксических моделей языка для систем обработки текстовой информации В. Н. Захаров, Ал-др А. Хорошилов, Ал-ей А. Хорошилов	4
Алгоритм обследования замкнутого помещения автономным мобильным роботом О. П. Архипов, А. В. Гасилов, Ю. А. Маньяков, О. А. Яковлев	10
Возможности преодоления защиты сети, использующей метаданные для управления сетевыми соединениями А. А. Грушо, Е. Е. Тимонина, С. Я. Шоргин	22
Модель гибридной интеллектуальной многоагентной системы гетерогенного мышления для информационной подготовки оперативных решений в региональных электрических сетях А. В. Колесников, С. В. Листопад	31
Визуальный язык моделирования эффекта синергии в гибридных интеллектуальных многоагентных системах И. А. Кириков, А. В. Колесников, С. В. Листопад	42
Априорное обратное гамма-распределение в байесовских моделях массового обслуживания А. А. Кудрявцев, С. И. Палионная, В. С. Шоргин	54
Интероперабельность как научно-методическая и нормативная основа бесшовной интеграции информационно-телекоммуникационных систем А. А. Башлыкова, А. А. Зацаринный, А. А. Каменщиков, С. В. Козлов, А. Я. Олейников, И. И. Чусов	61
Концептуальные подходы к созданию информационно-аналитического ситуационного центра в интересах стратегического планирования в области обеспечения национальной безопасности А. П. Сучков, А. В. Босов	73
Процессная модель модернизации и развития информационных систем на всех стадиях жизненного цикла А. П. Сучков	86

СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ

Том 28 № 4 Год 2018

СОДЕРЖАНИЕ

Некоторые системотехнические вопросы предоставления вычислительных ресурсов для научных исследований в гибридной высокопроизводительной облачной среде К. И. Волович	98
Основы живучести автоматизированных организаций информационного общества И. И. Быстров, В. Н. Веселов, К. К. Колин	110
Разработка приложения для подготовки расчетных сеток с градуирующими и изогнутыми краями для программной среды OpenFOAM Д. И. Читалов, С. Т. Калашников	122
Ситуационное планирование производства в сетевой M2M-системе А. В. Ильин, В. Д. Ильин	136
Элементы самообучения в системе извлечения биографических фактов T-парсер И. М. Адамович, О. И. Волков	145
Надкорпусная база данных коннекторов: развитие системы терминов проектирования И. М. Зацман, М. Г. Кружков	156
Метод описания структуры неоднословных коннекторов в надкорпусных базах данных О. Ю. Инькова, М. Г. Кружков	168
Тематические индикаторы переноса научных знаний в сферу технологий В. А. Минин, И. М. Зацман, В. А. Хавансков, С. К. Шубников	182
Об авторах	199
Авторский указатель за 2018 г.	203
2018 Author Index	210
Правила подготовки рукописей статей	217
Requirements for manuscripts	221

АВТОМАТИЧЕСКОЕ ПОСТРОЕНИЕ СИНТАКСИЧЕСКИХ МОДЕЛЕЙ ЯЗЫКА ДЛЯ СИСТЕМ ОБРАБОТКИ ТЕКСТОВОЙ ИНФОРМАЦИИ*

В. Н. Захаров¹, Ал-др А. Хорошилов², Ал-ей А. Хорошилов³

Аннотация: Описывается формальная модель синтаксической структуры текстов на основе обобщенных синтагм и методы ее автоматического построения. Предлагаемая синтаксическая модель базируется на машинной грамматике, в основу которой положена система флективных классов русских слов. Заложенное в теоретической концепции флективных классов слов русского языка жесткое соответствие между формой представления слов и их грамматической информацией позволило создать на этой основе новые классы — классы слов, имеющие одинаковые наборы грамматических признаков, соответствующие их формам представления в сходных контекстных окружениях.

Ключевые слова: синтаксическая модель языка; машинная грамматика; морфологический анализ; семантико-синтаксический анализ; формальная модель синтаксической структуры текста; обобщенные синтагмы

DOI: 10.14357/08696527180401

1 Введение

Последние несколько десятилетий задача обработки текстов на естественном языке не теряет своей актуальности. Направления исследований в этой области изменились незначительно, но методы трансформировались довольно существенно. Если раньше научные коллективы создавали инструментарий обработки текстов, используя экспертов-лингвистов для формирования правил семантико-синтаксического анализа, декларативных средств и др., то сейчас все большую популярность набирают методы машинного обучения, которые в перспективе должны заменить кропотливый ручной труд. К сожалению, на данный момент исследования в этой области еще не позволяют добиться стабильно высокого качества обработки.

*Работа выполнена при поддержке РФФИ (проект 18-37-00110 мол.а).

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

²Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, khoroshilov@mail.ru

³Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, a.a.horoshilov@mail.ru

Одной из центральных задач обработки текстов на естественном языке является формирование точных данных о синтаксической структуре текста. В настоящее время разработано значительное число синтаксических моделей текстов, все они являются условным отражением их структуры и ориентированы на решение конкретных практических задач автоматической обработки текстов. Любая синтаксическая модель всегда неполна и содержит ошибки, связанные с невозможностью установления однозначных правил синтаксического разбора текстов. В связи с этим возникает потребность в инструментарии, позволяющем постоянно пополнять синтаксические модели, опираясь на новые обучающие данные.

2 Синтаксическая модель языка на основе обобщенных синтагм

Концепция фразеологического концептуального анализа текстов, разработанная проф. Г. Г. Белоноговым совместно с его учениками, в рамках которой проводилось данное исследование, базируется на машинной грамматике, в основу которой положена система флективных классов русских слов [1]. Заложенное в теоретической концепции флективных классов слов русского языка жесткое соответствие между формой представления слов и их грамматической информацией позволило создать на этой основе новые классы — классы слов, имеющие одинаковые наборы грамматических признаков, соответствующие их формам представления в сходных контекстных окружениях.

При разработке этой синтаксической модели текстов авторы исходили из следующей гипотезы: одинаковым последовательностям обобщенных символов классов слов (обобщенным синтагмам) должны соответствовать одинаковые синтаксические структуры. При этом предполагалось, что такая гипотеза верна для любых синтаксических моделей и может быть полезна при решении как глобальных, так и частных задач синтаксического анализа.

В процессе реализации предлагаемой синтаксической модели была разработана система обобщенных символов классов слов, которая отражает систему грамматических категорий слов, а также грамматические признаки конкретной формы слова, согласованной с контекстным текстовым окружением. Было установлено, что номера флективных классов и грамматических окончаний позволяют однозначно установить все наборы основных грамматических признаков конкретной формы слова (род, число, падеж, лицо) [2]. Это сочетание двух параметров было выбрано в качестве элемента предлагаемой формальной модели синтаксической структуры текстов на основе обобщенных синтагм. Далее был сформирован двухбайтовый индекс обобщенных синтагм, состоящий из двух символов — символа номера флективного класса и символа его грамматического окончания. При назначении символов флективных классов, число которых составляет 231 класс, использовались буквенно-цифровые латинские и кириллические символы кодировочной таблицы 1251 (ANSI, WIN), при назначении символов грамматических окончаний, число которых составляет 72 окончания, также использовались буквенно-цифровые латинские и кириллические символы

Таблица 1 Фрагмент списка наименований понятий, полученных по тексту

Символы флективных классов слов	Символы грамматических окончаний слов
...	...
036 — <i>e</i>	<i>yt</i> — <i>o</i>
037 — <i>f</i>	<i>yю</i> — <i>p</i>
040 — <i>g</i>	<i>ы</i> — <i>q</i>
041 — <i>h</i>	<i>ые</i> — <i>r</i>
042 — <i>i</i>	<i>ый</i> — <i>s</i>
043 — <i>j</i>	<i>ым</i> — <i>t</i>
044 — <i>k</i>	<i>ыми</i> — <i>u</i>
045 — <i>l</i>	<i>ых</i> — <i>v</i>
...	...

кодирующей таблицы 1251. В табл. 1 приведены фрагменты списков буквенно-цифровых символов флективных классов и символов грамматических окончаний. Таким образом, синтаксическая модель языка будет выражена совокупностью устойчивых синтаксических конструкций, имеющих частоту встречаемости более трех. При этом корпус текстов, используемый для формирования синтаксической модели языка, должен быть значительного размера. Исследования проводились на корпусе текстов размером 1,5 млн предложений. В табл. 2 приведен фрагмент примера формирования последовательности обобщенных символов синтагм предложения. Здесь для каждого элемента предложения (по номеру флективного класса и буквенному коду грамматического окончания) был сформирован индекс обобщенной синтагмы слова. Как видно из этой таблицы, индекс в неявном виде содержит информацию о лексико-грамматическом классе слова, категории рода и одушевленности (для существительных), категории возвратности (для глаголов и отглагольных форм), принадлежности к краткой

Таблица 2 Фрагмент примера представления синтаксической структуры предложения в виде последовательности символов классов обобщенных синтагм

Слова исходного предложения	Грамматические признаки слов	Символ обобщенной синтагмы
...
резервн-ой	Полн. прил., ФК = 103 1) жен. р., ед. ч, род п. 2) жен. р., ед. ч, дат. п. 3) жен. р., ед. ч., тв. п. 4) жен. р., ед. ч., пред. п.	Фh
котельн-ой	Полн. субстантивированное прил., ФК = 103 1) жен. р., ед. ч., род. п. 2) жен. р., ед. ч., дат. п. 3) жен. р., ед. ч., тв. п. 4) жен. р., ед. ч., пред. п.	Фh
расположен-ы	Кратк. прич., ФК = 126, мн. ч.	жq
на	Предлог, ФК = 164, мод. упр. — 1) вин. п., 2) пред. п.	7A
...
Представление предложения в виде последовательности индексов обобщенных синтагм		
Aq Цf ЙЦ Фh Фh жq 7A Фi ЙS яA uA tS FB = Трубопроводы высокого давления резервной котельной расположены на значительном расстоянии от систем безопасности блока		

или полной форме прилагательного и причастия, типе его словоизменительной парадигмы, буквенном коде, длине словоизменительной основы и грамматического окончания. В нижней части таблицы показано представление предложения в виде последовательности индексов обобщенных синтагм. Эти последовательности формируют синтаксическую структуру предложения, элементами которой являются контактно расположенные объекты, обладающие грамматическими свойствами конкретных слов-эталонов. Конкретным элементам структуры могут соответствовать различные слова, грамматические признаки которых идентичны.

Необходимо отметить, что полученная синтаксическая структура предложения в виде последовательности символов обобщенных синтагм отражает полный набор морфологических и синтаксических признаков слов предложения; например, как видно, в этой структуре субстантивированное прилагательное «*котельная*» по грамматической форме является прилагательным, а по синтаксической функции в предложении выступает в роли существительного.

3 Алгоритм формирования синтаксической модели языка

В процессе исследований был разработан алгоритм, позволяющий автоматически формировать словарь устойчивых синтаксических конструкций для синтаксической модели языка. Далее приведем сам алгоритм.

- Шаг 1.** Разделить текст на предложения и выполнить обработку каждого анализируемого предложения процедурой морфологического анализа. Получить для каждого слова символ флективного класса и грамматического окончания.
- Шаг 2.** Провести синтаксический анализ предложения [3], провести назначение грамматической информации и построить дерево зависимости предложения. Рассчитать вероятности правильности получения информации.
- Шаг 3.** Выявить последовательности слов, для которых вероятность правильного установления грамматической информации равна 1.
- Шаг 4.** Сформировать для этих последовательностей индексы обобщенных синтагм.
- Шаг 5.** Сформировать из получившихся последовательностей подпоследовательности размером от трех слов до длины исходной последовательности.
- Шаг 6.** Добавить подпоследовательности индексов обобщенных синтагм, полученные на шаге 5, в словарь устойчивых синтаксических конструкций. Пересчитать значения частот для этих подпоследовательностей.
- Шаг 7.** Исключить из словаря устойчивых синтаксических конструкций синтагмы с частотой менее трех.

4 Заключение

В статье описана модель синтаксической структуры текстов на основе обобщенных синтагм и методы ее автоматического формирования. Предлагаемая синтаксическая модель базируется на машинной грамматике, в основу которой положена система флективных классов русских слов. В процессе проведения исследований была сформирована синтаксическая модель для русского языка, исходными данными для ее создания служил массив, содержащий 1,5 млн предложений. На выходе был получен словарь устойчивых синтаксических конструкций размером более 40 тыс. Данная модель может быть использована для многих задач обработки текстовой информации, среди которых задачи машинного перевода, синтаксического анализа текстов, выявления понятийного состава текстов, разрешения омонимии и других. Достоинством данной модели является простота обучения и возможность ее использования для любого естественного языка.

Литература

1. Белоногов Г. Г. Теоретические проблемы информатики. Т. 2: Семантические проблемы информатики / Под общ. ред. К. И. Курбакова. — М.: РЭА им. Г. В. Плеханова, 2008. 342 с.
2. Аблов И. В., Козичев В. Н., Ширманов А. В., Хорошилов Ал-др А., Хорошилов Ал-ей А. Средства машинной грамматики русского языка (по Г. Г. Белоногову) // Научно-техническая информация. Сер. 2. Информационные процессы и системы, 2018. № 6. С. 32–46.
3. Захаров В. Н., Хорошилов А. А. Автоматическое формирование визуального представления смыслового содержания документа // Системы и средства информатики, 2013. Т. 23. № 1. С. 143–158.

Поступила в редакцию 10.09.18

AUTOMATIC CONSTRUCTION OF SYNTACTIC LANGUAGE MODELS FOR TEXT PROCESSING SYSTEMS

V. N. Zakharov¹, Al-dr A. Khoroshilov², and Al-ey A. Khoroshilov²

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

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

Abstract: The article describes a formal model of the syntactic structure of texts based on generalized syntagmas and methods of its automatic construction. The proposed syntactic model is based on machine grammar which is based on the

system of inflectional classes of Russian words. The correspondence between the form of the representation of words and their grammatical information, laid down in the theoretical conception of the inflectional classes of words in the Russian language, made it possible to create on this basis new classes — classes of words that have the same sets of grammatical features corresponding to their forms of representation in similar context.

Keywords: syntactic model of language; computer grammar; morphological analysis; semantic and syntactic analysis; formal model of syntactic text structure; generalized syntagmas

DOI: 10.14357/08696527180401

Acknowledgments

The research was supported by the Russian Foundation for Basic Research (project 18-37-00110 mol_a).

References

1. Belonogov, G. G. 2008. *Teoreticheskie problemy informatiki. Semanticheskie problemy informatiki* [Theoretical problems of informatics. Semantic problems of informatics]. Moscow: G. V. Plekhanova REA. 342 p.
2. Ablov, I. V., V. N. Kozichev, A. V. Shirmanov, Al-dr A. Khoroshilov, and Al-ey A. Khoroshilov. 2018. The tools of a machine grammar of the Russian language (based on G. G. Belonogov). *Automatic Documentation Math. Linguistics* 52(3):142–156.
3. Zakharov, V. N., and A. A. Khoroshilov. 2013. Avtomaticheskoe formirovanie vizual'nogo predstavleniya smyslovogo sodержaniya dokumenta [Automatic generation of visual representation of the document's semantic content]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 13(1):143–158.

Received September 10, 2018

Contributors

Zakharov Victor N. (b. 1948) — Doctor of Science in technology, associate professor, scientific secretary, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow, Russian Federation; vzakharov@ipiran.ru

Khoroshilov Aleksandr A. (b. 1952) — Doctor of Science 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, Russian Federation; khoroshilov@mail.ru

Khoroshilov Alexey A. (b. 1988) — Candidate of Science (PhD) in technology, scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow, Russian Federation; a.a.horoshilov@mail.ru

АЛГОРИТМ ОБСЛЕДОВАНИЯ ЗАМКНУТОГО ПОМЕЩЕНИЯ АВТОНОМНЫМ МОБИЛЬНЫМ РОБОТОМ

О. П. Архипов¹, А. В. Гасилов², Ю. А. Маньяков³, О. А. Яковлев⁴

Аннотация: Предлагается алгоритм обследования замкнутого помещения автономным мобильным роботом, оснащенный регистрирующими устройствами. Осуществлена постановка задачи обследования. Описанный алгоритм использует разработанную модель представления плана помещения, основанную на представлении замкнутого помещения в виде невыпуклого многоугольника без самопересечений, с дырками, который итеративно формируется на основе многоакурсной реконструкции, в результате чего осуществляется его уточнение и наращивание по мере обследования помещения. Кроме того, данная модель позволяет формализовать условие полной обследованности замкнутого помещения, которое может быть проверено за ограниченное время, и позволяет построить граф для поиска пути и обхода препятствий. Описаны основные этапы алгоритма, проведен анализ его вычислительной сложности, а также представлены результаты работы алгоритма на основе синтетических наборов входных данных. Рассмотрен подход к решению задачи построения маршрута обследования, а также приведено описание алгоритма объединения многоугольников плана помещения.

Ключевые слова: компьютерное зрение; автономный мобильный робот; план помещения; навигация; поиск пути

DOI: 10.14357/08696527180402

1 Введение

Одной из сфер применения трехмерной реконструкции является построение трехмерных моделей помещений, внутренних пространств геологических образований и других подобных объектов. В случае отсутствия возможности доступа человека либо при условии высокой сложности и больших размеров данных областей наиболее актуальным сценарием является использование автономного агента обследования, например автономной мобильной роботизированной системы, оснащенной необходимым для реконструкции оборудованием.

¹Орловский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, arkhprov12@yandex.ru

²Орловский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, gasilov.av@ya.ru

³Орловский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, maniakov_yuri@mail.ru

⁴Орловский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, maucra@gmail.com

В рамках реализации решения глобальной задачи обследования помещений автономным агентом возникает задача построения маршрута обследования. Формально эта задача может быть сформулирована следующим образом.

Пусть есть замкнутое помещение и некоторый агент обследования, находящийся в некоторой его точке. Агент характеризуется следующими элементами: способен отслеживать свое положение в помещении; обладает регистрирующим устройством, позволяющим получать информацию об удаленности различных объектов от него (стереокамера, камера глубины, LIDAR (light identification detection and ranging) и т. п.); способен поворачиваться на 360° в горизонтальной плоскости и перемещаться по прямой траектории.

Задача построения маршрута обследования помещения состоит в получении такого маршрута перемещения субъекта обследования, после прохождения которого в помещении не останется необследованных зон.

В данной работе предлагается алгоритм построения такого маршрута обследования, а также модель хранения плана помещения, строящаяся на основе исходных данных в виде воксельной модели.

Согласно работам [1, 2], существующие модели хранения операционной среды агента можно разделить на три класса: метрическая, координатная; топологическая, сетевая; семантическая, предикатная.

Метрические модели обычно направлены на использование в статических средах: они в явном виде хранят координаты всех препятствий. Данный подход требует больше памяти по сравнению с другими, однако позволяет позиционировать агента с некоторой определенной точностью. Траектория движения в таких моделях представляет собой последовательность координат, которые необходимо посетить агенту для достижения своей цели [2].

Топологические модели представляют операционную среду в виде графа, где вершины обозначают зоны окружающей среды, а ребра определяют их непосредственную достижимость друг из друга. Однако особенностью такого подхода является то, что вершины обозначают зоны нефиксированного размера и даже не обязаны хранить информацию о проходимости внутри зоны. Для решения этой проблемы применяется оперативный анализ окружающей обстановки, который позволяет агенту находить вход и выход из зоны [1].

Семантические модели описывают операционную среду в виде набора предикатных выражений. Такой подход позволяет обрабатывать и приводить к общему виду информацию, поступающую с датчиков различных типов. В основном такие модели предназначены для вспомогательных целей (например, интерпретация команды на естественном языке, анализ окружающей среды на наличие особых фактов) и не используются для прокладывания путей [2].

Предлагаемая модель хранения данных согласно приведенной выше классификации относится к категории метрических. При этом она может использоваться как для описания операционной среды агента, так и для получения человекочитаемого плана помещения.

2 Алгоритм обследования помещения

В процессе функционирования агент работает с двумя планами помещения. Первый — план текущей видимой части помещения, второй — накопленный план, составленный из объединения всех полученных на каждой итерации частей.

На основе данных действий агент управляется с помощью следующего алгоритма: получить информацию об окружении, обновить накопленный план полученной информацией, проверить достижение условия остановки и либо успешно завершить обследование помещения, либо переместиться в новую точку для получения дополнительной информации.

Блок-схема алгоритма представлена на рис. 1.

Информация о видимости получается в виде воксельной модели [3], строящейся на основе облака точек, попадающего в область видимости, полученного с помощью оптических датчиков, сканеров глубины или иных подобных датчиков.

В рамках данной работы рассматривается агент с полным круговым обзором с бесконечным разрешением, однако все приведенные положения справедливы и применимы для агента с ограниченным углом обзора и конечным разрешением.

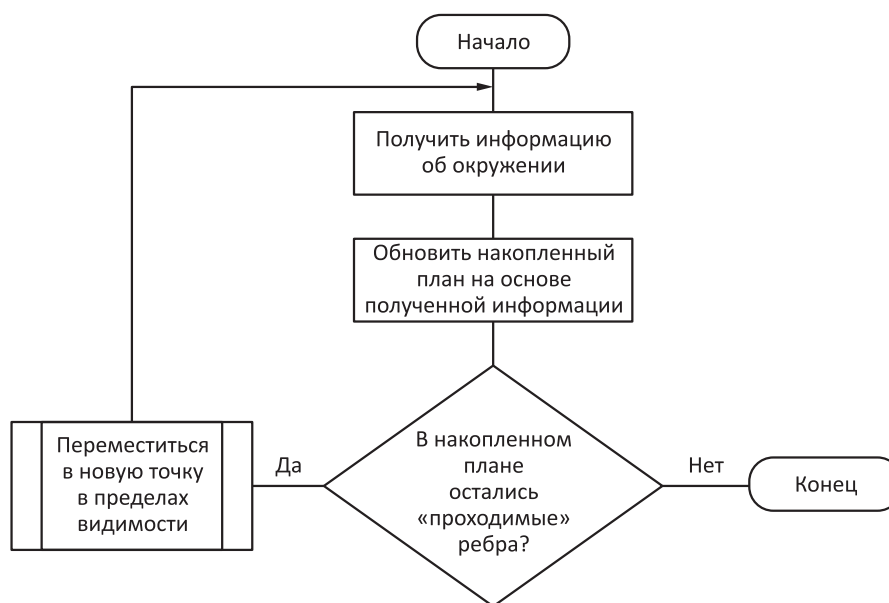


Рис. 1 Блок-схема алгоритма обследования помещения

3 Модель хранения плана помещения

В качестве модели предлагается хранить план в виде единственного невыпуклого многоугольника с дырками, без самопересечений (рис. 2). При этом каждому ребру присваивается одна из двух меток: либо «непроходимое ребро» (для стен, препятствий), либо «проходимое ребро» (для границ видимости).

Существенные преимущества данной модели — простота отслеживания границ видимости для вычисления направления движения агента и простота формализации условия завершения обследования помещения: помещение является обследованным, если в плане, описывающем его, не осталось ни одного ребра с меткой «проходимое ребро».

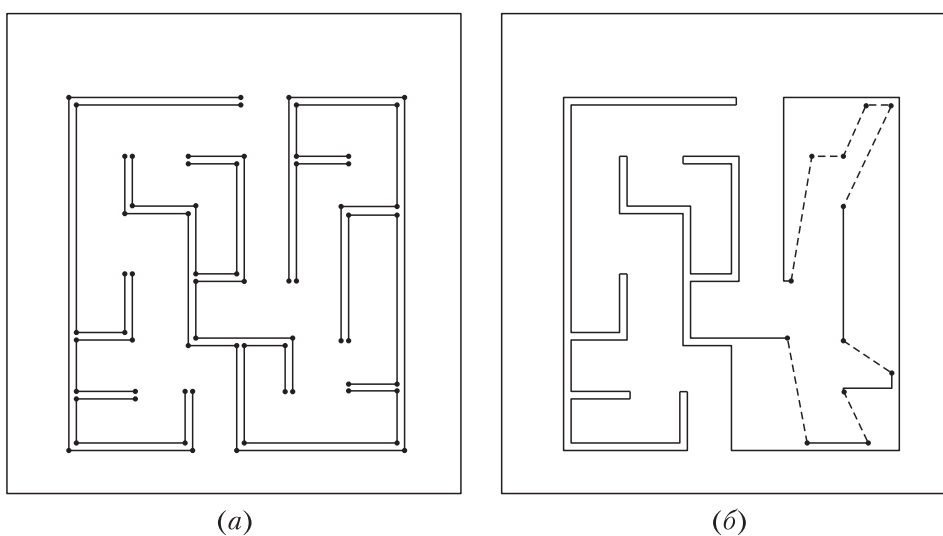


Рис. 2 Пример помещения, представленного с помощью предлагаемой модели: (а) модель помещения; (б) модель частично обследованного помещения (непрерывные линии — непроходимые ребра, прерывистые линии — проходимые ребра)

3.1 Получение информации об окружении

Агент получает информацию об окружении в виде воксельной модели, которая однозначно соотносится с имеющимся планом на основании известного текущего положения агента.

Так как планы хранятся в виде многоугольника, необходимо привести воксельную модель к формату многоугольника. Предлагается спроецировать модель на плоскость $Z = 0$. В результате данной операции получается модель, состоящая из множества квадратов, к которой потом применяется следующая операция: ви-

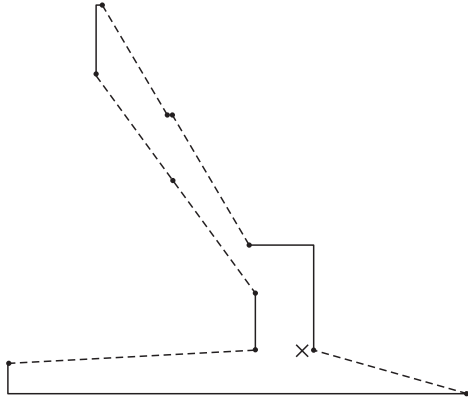


Рис. 3 Пример полученной информации об окружении. Многоугольник построен по всем точкам, видимым из текущей, а ребра помечены в соответствии с условием: непрерывные линии — непроходимые ребра, прерывистые линии — проходимые ребра, крестик — текущее положение агента

димые из положения агента ребра квадратов сортируются по углу относительно текущей позиции наблюдения и в данном порядке добавляются в новый многоугольник текущей видимости.

После этого происходит маркировка ребер: если $\text{dist}(v_i, v_j) < \alpha$, где $\text{dist}(x, y)$ — евклидово расстояние между двумя точками на плоскости; v_i и v_j — координаты двух последовательных вершин в полученном многоугольнике, то считается, что воксели, соответствующие данным точкам, расположены достаточно близко, чтобы агент не мог пройти между ними, и такое ребро получает метку «непроходимое ребро». В случае если данное условие не выполняется, ребро получает метку «проходимое ребро» (рис. 3).

При размере агента больше воксельной ячейки может возникнуть проблема недостижимых проходимых ребер, пример которой приведен на рис. 4: агент через узкий проход видит объекты, из-за которых, несмотря на непосредственную недостижимость этих объектов, образуются недостижимые проходимые ребра.

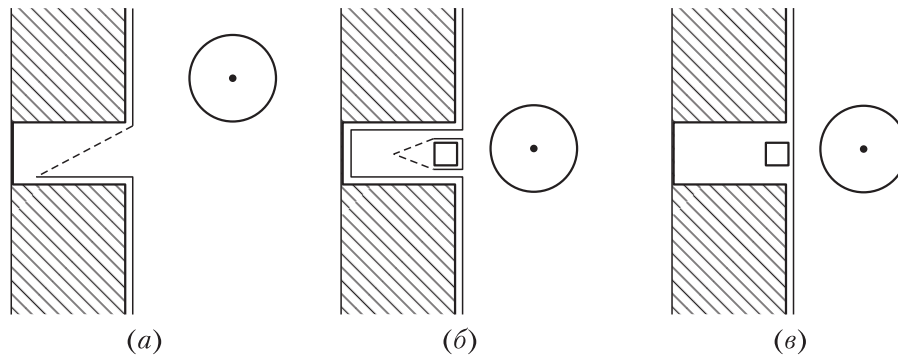


Рис. 4 Пример проблемы недостижимых проходимых ребер: (а) пример проблемы, разрешимой с помощью алгоритма построения маршрута; (б) пример проблемы, не разрешимой с помощью алгоритма построения маршрута; (в) пример проблемы, решенной предлагаемым методом

В некоторых случаях (рис. 4, а) будет возможно увидеть это ребро из какой-либо точки, не достигая его непосредственно, и таким образом убрать его, однако возможны случаи (рис. 4, б), когда эти ребра не будут обследованными из любого возможного положения агента, что ведет к невозможности достижения условия завершения работы алгоритма.

Данную проблему удобно решать на этапе построения многоугольника текущей видимости. Подход к ее решению заключается в проверке расстояния между всеми вершинами в текущем многоугольнике видимости. И в случае если две несоседних вершины удовлетворяют условию $\text{dist}(v_i, v_j) < \alpha$, необходимо убрать из многоугольника текущей видимости все вершины, лежащие между ними в порядке обхода, и добавить непроходимое ребро от v_i до v_j . При этом из двух возможных интервалов выбирается интервал с минимальным углом.

Значение α целесообразно выбирать, опираясь на реальные габариты агента. В рамках прототипа форма агента описывается окружностью диаметром S . Значение S принималось равным единице, а значение α — равным $\sqrt{2}$.

3.2 Объединение планов

Необходимой операцией для работы с данной моделью является объединение двух планов: текущего накопленного плана и новой поступившей информации.

Задача построения объединения также известна как задача о построении оверлеев многоугольников и имеет несколько решений, рассмотренных в [4]. В рамках реализации прототипа был применен метод построения оверлеев с помощью триангуляции [5] как наиболее быстрый алгоритм, способный корректно работать на многоугольниках заданного типа.

Выбранный алгоритм объединения многоугольников состоит из следующих этапов.

1. Построение триангуляции с ограничениями по ребрам двух входных многоугольников.
2. Классификация каждого треугольника триангуляции по принадлежности исходным многоугольникам.
3. Классификация каждого треугольника триангуляции по принадлежности результирующему многоугольнику на основе выбранной операции (в данном случае — объединение) и данных предыдущего этапа.
4. Все треугольники, помеченные принадлежащими результирующему многоугольнику, объединяются в один многоугольник, который и будет служить результатом.

Важным моментом при объединении планов является корректная обработка меток ребер: при пересечении коллинеарных ребер с различными метками результирующее ребро будет помечено как «непроходимое», а в остальных случаях метки исходных ребер сохраняются.

4 Построение маршрута обследования

В рамках предложенной модели хранения и обновления плана допустимы различные подходы к реализации алгоритма прокладывания маршрута, однако каждая реализация должна следовать правилу: всегда двигаться с целью достижения «проходимых» ребер плана.

Если следовать такому правилу, то, исходя из способа хранения и обновления плана, а также того, что согласно условию ведется обследование замкнутого помещения, в какой-то момент будет достигнута ситуация, когда в текущем плане обследованного помещения не останется ребер с меткой «проходимое ребро», и, следовательно, цель алгоритма обследования помещения будет достигнута.

В рамках разработанного прототипа агент использует следующий алгоритм выбора пути: следовать в ближайшее по абсолютному расстоянию проходимое ребро кратчайшим путем. Блок-схема алгоритма приведена на рис. 5.

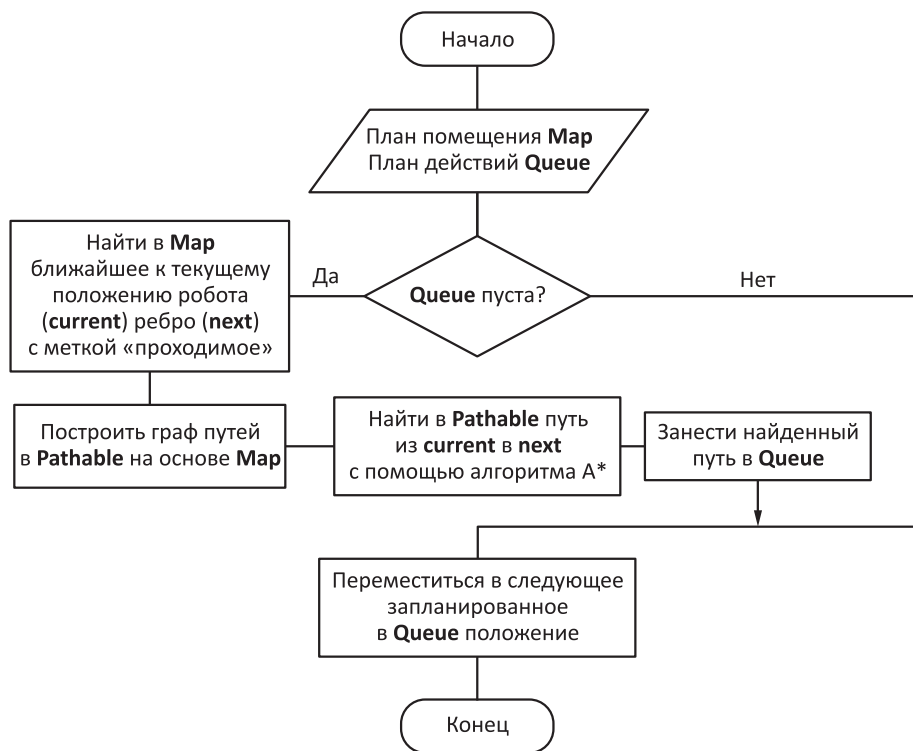


Рис. 5 Блок-схема алгоритма перемещения агента

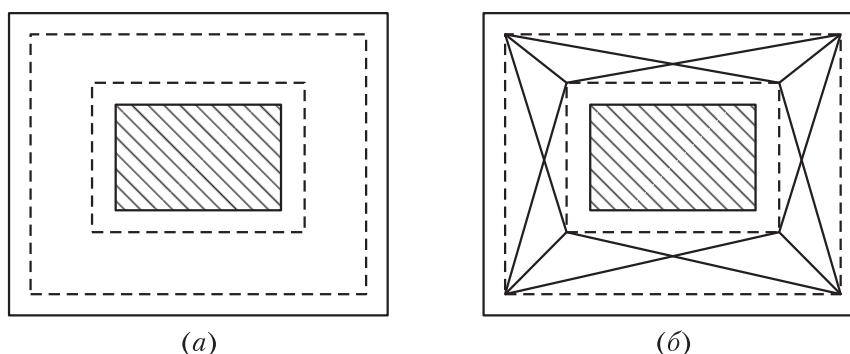


Рис. 6 Пример плана со смещенными вершинами (непрерывная линия — оригинальный план; прерывистая линия — план со смещенными вершинами) (а) и пример графа достижимости (б)

Для обхода препятствий и поиска кратчайшего пути к выбранной точке используется алгоритм A^* [6] для поиска пути в специальном графе достижимости, построенном по известному накопленному плану.

Для упрощения задачи и построения такого графа необходимо привести задачу к $S = 0$ путем добавления дополнительной граничной области шириной $d = S/2$ к контуру каждого препятствия (рис. 6, а). Непосредственно сам граф достижимости строится по накопленному плану помещения: вершинам графа соответствуют вершины многоугольника, а ребро между вершинами существует, если соответствующие вершины многоугольника соединены ребром или видимы друг из друга (рис. 6, б).

5 Анализ эффективности

В рамках предлагаемой модели и ее применения отмечаются три основных этапа: получение информации об окружении; слияние планов; построение маршрута.

На этапе получения информации об окружении входными данными служит список видимых в данный момент точек, и оценка его сложности выходит за рамки статьи. Однако обработка входных данных также важна. Операция первичной разметки ребер требует однократного прохода по всем вершинам многоугольника с проверкой заданного условия, ее вычислительная сложность составляет $O(n)$, где n — число вершин в многоугольнике текущей видимости. Одно из возможных решений второй операции, направленной на исключение недостижимых проходимых ребер, требует проверки расстояний между всеми парами вершин в многоугольнике текущей видимости с последующей их обработкой, что дает вычислительную сложность $O(n^2)$.

Операция слияния планов имеет ряд возможных решений различной сложности реализации и различной вычислительной сложности [4]. Реализованный в прототипе алгоритм построения оверлеев на основе триангуляции в среднем случае имеет вычислительную сложность $O(n)$ и $O(n \log(n_1 + n_2))$ в худшем случае, где n_1 — число вершин первого многоугольника; n_2 — число вершин второго многоугольника; n_0 — число вершин на пересечении ребер двух многоугольников; $n = n_1 + n_2 + n_0$.

Этап построения текущего маршрута состоит из двух основных операций: построения графа и поиска пути по нему. В предлагаемой реализации этапа построения маршрута описанная операция построения графа потребует $O(n)$ операций на смещение всех вершин многоугольника (где n — число вершин такого многоугольника) и $O(n^2)$ операций на получение возможных путей между всеми парами вершин. Если принять во внимание инкрементальную природу процесса обследования, то имеется возможность не строить на каждой итерации новый граф, а перестраивать граф, полученный на предыдущей итерации, добавляя каждую новую вершину за $O(n^2)$.

Также на некоторых этапах возможно проведение эвристической оптимизации над получаемыми многоугольниками: объединять соседние ребра, имеющие одинаковые метки, в случае их коллинеарности. На тестовых моделях такая оптимизация позволила снизить число ребер в среднем в 10 раз.

6 Результаты

Описанная модель и подход к ее применению были реализованы в рамках программного прототипа. Данный прототип включает в себя эмулятор, реализующий функции получения информации о видимой части помещения и функции перемещения для заданного входного файла помещения.

Работоспособность модели и корректность достижения поставленной цели были проверены на ряде синтетических тестов. Пример промежуточных шагов работы алгоритма обследования помещения представлен на рис. 7.

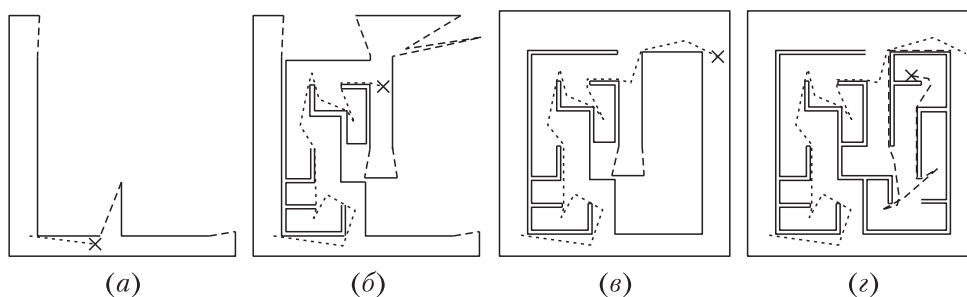


Рис. 7 Пример промежуточных шагов работы алгоритма

7 Выводы

В рамках данной статьи был предложен алгоритм построения маршрута обследования помещения, а также модель хранения и обработки плана помещения для решения задачи обследования помещения на основе многоугольника специального вида. Были рассмотрены возможные проблемы, а также проведена оценка вычислительной сложности предлагаемого подхода. Предлагаемая модель хранения плана не имеет принципиальных требований к специфике задачи и может применяться в рамках любой задачи обследования с агентом, удовлетворяющим заданным минимальным требованиям.

Литература

1. *Siegwart R., Nourbakhsh I. R.* Introduction to autonomous mobile robots. — Cambridge, MA, USA: MIT Press, 2004. 321 p.
2. *Давыдов О. И., Платонов А. К.* Сеть Пассфреймов — комбинированная модель операционной среды мобильного робота // Препринты ИПМ им. М. В. Келдыша, 2015. № 15. 28 с. <http://library.keldysh.ru/preprint.asp?id=2015-15>.
3. *Яковлев О. А.* Методы инициализации воксельного объема в задаче трехмерной реконструкции // Системы и средства информатики, 2018. Т. 28. № 1. С. 53–64.
4. *Ченцов О. В., Скворцов А. В.* Обзор алгоритмов построения оверлеев многоугольников // Вестник Томского государственного университета, 2003. № 280. С. 338–345.
5. *Скворцов А. В.* Построение объединения, пересечения и разности произвольных многоугольников в среднем за линейное время с помощью триангуляции // Вычислительные методы и программирование: новые вычислительные технологии, 2002. Т. 3. № 1. С. 116–123.
6. *Hart P. E., Nilsson N. J., Raphael B.* A formal basis for the heuristic determination of minimum cost paths // IEEE T. Syst. Sci. Cyb., 1968. Vol. 4. Iss. 2. P. 100–107.

Поступила в редакцию 09.08.18

ENCLOSED ROOM EXPLORATION ALGORITHM FOR AN AUTONOMOUS MOBILE ROBOT

O. Arkhipov, A. Gasilov, Yu. Maniakov, and O. Yakovlev

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

Abstract: An enclosed room exploration algorithm for an autonomous mobile robot equipped with sensors is introduced. This paper includes the definition of the task of enclosed room exploration. The algorithm utilizes the data model

of a floorplan, based on a concave polygon with holes, which is refined and enhanced by merging different views of the room in the iterative manner during exploration. This representation allows formulating the termination criterion that can be checked in constant time and constructing a graph for pathfinding and obstacle avoidance purposes. In addition, analysis of algorithm's computational complexity and results of a test based on synthetic datasets are provided. Furthermore, an approach to exploration path building and an algorithm of floorplan's polygons union are considered.

Keywords: computer vision; autonomous exploration mobile robot; floorplan; navigation; pathfinding

DOI: 10.14357/08696527180402

References

1. Siegwart, R., and I. R. Nourbakhsh. 2004. *Introduction to autonomous mobile robots*. Cambridge, MA: MIT Press. 321 p.
2. Davydov, O. I., and A. K. Platonov. 2015. Set' Passfreymov — kombinirovannaya model' operatsionnoy sredy mobil'nogo robota [Passframe Network — combined operating environment model for a mobile robot]. *Preprinty IPM im. M. V. Keldysha* [Keldysh Institute preprints]. 28 p. Available at: <http://library.keldysh.ru/preprint.asp?id=2015-15> (accessed July 31, 2018).
3. Yakovlev, O. A. 2018. Metody initsializatsii voksel'nogo ob"ema v zadache trekhmernoy rekonstruktsii [Initial bounding box estimation methods for volumetric three-dimensional reconstruction]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(1):53–64.
4. Chentsov, O. V., and A. V. Skvortsov. 2003. Obzor algoritmov postroeniya overleev mnogougol'nikov [A review of the algorithms of polygon overlays design]. *Tomsk State University J.* 280:338–345.
5. Skvortsov, A. V. 2002. Postroenie ob"edineniya, peresecheniya i raznosti proizvol'nykh mnogougol'nikov v srednem za lineynoe vremya s pomoshch'yu triangulyatsii [An algorithm for constructing the union, intersection, and difference of arbitrary polygons on the basis of triangulation with linear-time complexity on average] *Numerical Methods and Programming* 3(1):116–123.
6. Hart, P. E., N. J. Nilsson, and B. Raphael. 1968. A formal basis for the heuristic determination of minimum cost paths. *IEEE T. Syst. Sci. Cyb.* 4(2):100–107.

Received August 9, 2018

Contributors

Arkhipov Oleg P. (b. 1984) — director, Orel Branch of the Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences; 137 Moskovskoe Shosse, Orel 302025, Russian Federation; arkhipov12@yandex.ru

Gasilov Artur A. (b. 1992) — junior scientist, Orel Branch of the Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences; 137 Moskovskoe Shosse, Orel 302025, Russian Federation; gasilov.av@ya.ru

Maniakov Yuri A. (b. 1984) — senior scientist, Orel Branch of the Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences; 137 Moskovskoe Shosse, Orel 302025, Russian Federation; maniakov_yuri@mail.ru

Yakovlev Oleg A. (b. 1992) — junior scientist, Orel Branch of the Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences; 137 Moskovskoe Shosse, Orel 302025, Russian Federation; maucra@gmail.com

ВОЗМОЖНОСТИ ПРЕОДОЛЕНИЯ ЗАЩИТЫ СЕТИ, ИСПОЛЬЗУЮЩЕЙ МЕТАДААННЫЕ ДЛЯ УПРАВЛЕНИЯ СЕТЕВЫМИ СОЕДИНЕНИЯМИ*

А. А. Грушо¹, Е. Е. Тимонина², С. Я. Шоргин³

Аннотация: Управление соединениями в сети с помощью метаданных не рассматривает время реализации нескольких информационных технологий. Этот факт может позволить преодолевать защиту сети, использующей метаданные для управления сетевыми соединениями. Для решения потенциальных проблем, связанных с этой уязвимостью, предложено внесение дополнений в функцию управления с помощью метаданных. В частности, предложен метод построения очередей и метод принудительной реконфигурации сети. В случае SDN (Software Defined Network) метаданные непосредственно определяют для контроллера момент времени и требования к реконфигурации сети.

Ключевые слова: информационная безопасность; информационные технологии; распределенные информационно-вычислительные системы; метаданные; программно-конфигурируемые сети; составные задачи

DOI: 10.14357/08696527180403

1 Введение

Управление соединениями хостов в сети с помощью метаданных [1, 2] решает ряд важных вопросов обеспечения информационной безопасности в распределенных информационно-вычислительных системах (РИВС). Если РИВС поддерживается SDN, то метаданные обеспечивают дополнительную изоляцию системы управления сетью и бизнес-процессов от информационных потоков в сети.

Однако управление соединениями в сети с помощью метаданных не рассматривает ряд параметров функционирования сети. Одним из таких параметров является время реализации нескольких информационных технологий. Этот факт может позволить преодолевать защиту сети, использующей метаданные для управления сетевыми соединениями.

Для решения потенциальных проблем, связанных с этой уязвимостью, предложено внесение дополнений в функцию управления с помощью метаданных.

*Работа поддержана РФФИ (проект 16-11-10227).

¹Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, grusho@yandex.ru

²Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, eltimon@yandex.ru

³Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, sshorgin@ipiran.ru

В частности, предложен метод построения очередей и метод принудительной реконфигурации сети. В случае SDN метаданные непосредственно определяют для контроллера момент времени и требования к реконфигурации сети.

2 Модель метаданных

Работа, выполняемая РИВС, сводится к решению задач, реализуемыми приложениями [3]. Решение задач состоит из трех процессов:

- (1) сбора информации для решения задачи (исходные данные);
- (2) обработки информации на компьютерах с помощью программного обеспечения (приложений);
- (3) распределения результатов обработки информации.

Информационные технологии можно представить в виде составных задач [4], поэтому можно объединять понятия информационной технологии и задач.

Компьютер как узел сети называется хостом. Различные задачи можно решать на различных хостах сети. Тогда сеть позволяет собирать исходные данные для задач и распределять результат обработки. Следовательно, существует бинарное отношение $H(A)$ между хостом H и конкретным экземпляром задачи A , представленной программным обеспечением и информационными ресурсами.

Политика безопасности РИВС требует контроля за взаимодействиями хостов в сети, который сводится к мониторингу взаимодействий и управлению соединениями. Контроль взаимодействий хостов в сети позволяет снизить угрозы внедрения и распространения вредоносного кода через сетевое оборудование и каналы связи. В работах [1, 2] управление взаимодействиями хостов в сети предлагается реализовать с помощью метаданных.

Обозначим множество задач, реализуемых в РИВС, через Ω и рассмотрим на нем бинарное отношение порождения (A_i, A_j) , где A_i и A_j — задачи, причем A_i определяет запуск или участвует в запуске задачи A_j . Порождения, в которых исходная задача находится на одном хосте, а порождаемая — на другом, обозначим через \mathcal{B} и будем называть *метаданными*.

Для задач, входящих в отношение \mathcal{B} , определим три дополнительные задачи \mathcal{M} , \mathcal{N} и \mathcal{R} , которые управляют взаимодействиями в сети на основе метаданных \mathcal{B} . Задача \mathcal{M} распределяет приложения для решения задач между хостами, для простоты будем говорить о распределении задач на хостах. Задача \mathcal{M} определяет бинарное отношение $H(A)$, означающее, что на хосте H может вычисляться задача A .

Результаты задачи \mathcal{M} используются задачей \mathcal{N} . Задача \mathcal{N} поддерживает связь с каждым хостом и отвечает за разрешение и предоставление хостам информации по запросу о взаимодействиях задач на разных хостах. Разрешение основывается на метаданных \mathcal{B} .

Задача \mathcal{R} строит основной и резервные маршруты по заданию задачи \mathcal{N} . Например, \mathcal{R} находится в контроллере сети SDN.

Пусть на хосте $H(A)$ легально запущена задача A . Задача A_1 находится в отношении порождения (A, A_1) и расположена на другом хосте. На каждом хосте H есть агент с криптографическими средствами и ключом $k(H)$ для связи с хостом $H(\mathfrak{N})$, причем для каждого H соединение с $H(\mathfrak{N})$ восстанавливается быстро.

Для обращения к задаче A_1 задача A через агента хоста $H(A)$ связывается с задачей \mathfrak{N} , которая определяет наличие порождения (A, A_1) . Тогда на хост $H(A_1)$ через агента этого хоста направляются информация о необходимости соединения с $H(A)$, ключ $k(A, A_1)$ для защиты этого соединения, идентификатор, порт и время. Аналогичная информация направляется на хост $H(A)$. После выполнения задачи A_1 задача A получает результаты, а соединение $H(A)$ с $H(A_1)$ разрывается.

3 Возможности преодоления защиты сети, использующей метаданные

В концепции сетевой защиты на основе метаданных, определенной в работах [1–4], нет упоминания о временных характеристиках. Напомним, что метаданные содержат информацию о допустимости соединения хостов $H(B)$ и $H(A)$ на основе необходимости взаимодействия задач B и A в реализуемой информационной технологии.

Вместе с тем в информационной системе могут одновременно функционировать несколько информационных технологий, которые имеют право обращаться к задаче A . Тогда возможно возникновение очереди за использованием задачи A .

Предположим, что в ходе выполнения технологий образовалась очередь из задач B_1, B_2, \dots, B_s . Распределение задач по хостам осуществляется на уровне моделей бизнес-процессов с помощью задачи \mathfrak{M} , и это распределение задач можно считать случайным. Запуск нескольких информационных технологий и решение последовательности задач, реализующих эти технологии, можно считать случайным процессом.

Если рассмотреть конкретную задачу A и ее хост $H(A)$, то в ходе выполнения нескольких информационных технологий можно полагать, что обращения к задаче A являются независимыми случайными событиями. Поскольку метаданные не контролируют размер очереди, то возможно нарушение информационной безопасности за счет превышения возможностей хоста $H(A)$ устанавливать соединения с хостами $H(B_1), H(B_2), \dots, H(B_s)$ задач B_1, B_2, \dots, B_s , ожидающих обращения к задаче A .

Данное нарушение защищенности сходно с атакой типа DDoS (Distributed Denial of Service). Различие состоит в том, что недопустимая очередь может возникнуть случайно в различных хостах сети.

Предлагается следующее решение. Задача \mathfrak{N} получает от задачи A через агента в хосте $H(A)$ сообщение о том, что A занята. Агент в $H(\mathfrak{N})$ сообщает об этом всем хостам $H(B_1), H(B_2), \dots, H(B_s)$, подавшим заявки в задачу \mathfrak{N} о соединении с задачей A . В этом случае задача \mathfrak{N} может сформировать

очередь, последовательно сообщая хостам $H(B_i)$, $i = 1, \dots, s$, об их возможности соединения с хостом $H(A)$, т.е. хост $H(A)$ должен посылать хосту $H(\mathfrak{N})$ сообщение об освободившейся возможности решения задачи A .

Такой подход обеспечивает разрешение конфликта между информационными технологиями и не приводит к сбою типа DDoS. Однако время ожидания продолжения информационных технологий становится неопределенным и случайным. Кроме того, этот метод значительно увеличивает нагрузку на задачу \mathfrak{N} . Поэтому этот метод надо «усилить» возможностью реконфигурации сети и модернизации метаданных [5].

4 Вероятностные модели

Пусть время решения экземпляра задачи A экспоненциально с параметром μ . Если получены некорректные данные для задачи A , то ее анализ и решение могут затянуться. Это влечет за собой срыв сроков исполнения всех задач в очереди и задержку в завершении нескольких информационных технологий. Поэтому при задержке решения задачи необходимо по возможности реконфигурировать распределение задач по хостам. Например, можно выделить хост с новым экземпляром задачи A и перенаправить очередь на хост $\hat{H}(A)$ (при необходимости заново стартуя циклы). Таким образом определяется запрет на продолжение действующего вычислительного процесса, который связан с разладкой вероятностного распределения обработки данных. Теория запретов вероятностных мер в конечных пространствах рассматривалась в работах [6–8].

В сделанных предположениях о распределении стандартного решения задачи A вероятность превышения временного порога T_0 равна $e^{-\mu T_0}$. При превышении порога T_0 начинается процедура реконфигурации сети. Поскольку установка нового экземпляра задачи A и реконфигурация сети — это сложные задачи, то возникает задача оптимизации между допустимым временем ожидания и началом реконфигурации.

Второй задачей, приводящей к необходимости реконфигурации, является появление очереди недопустимо большой длины для хотя бы одной из задач информационной системы. Здесь также возникает задача оптимизации, которая опирается на вероятностное моделирование семейства очередей. При этом требуется, чтобы реконфигурация сети для обхода длинных очередей минимально влияла на работу информационной системы.

В основу построения таких вероятностных моделей положим тот факт [9], что распределенная система может быть представлена в виде одного компьютера, в котором единый планировщик выстраивает последовательность текущих решаемых задач (независимо от числа ядер и процессоров). В этой модели не важно, какие задачи обращаются к задаче A , а важно число экземпляров A , необходимых для удовлетворения запросов очереди. Тогда последовательность задач на обработку описывает ситуацию во всей информационной системе для

всех запрашиваемых задач. Например, не важно, что к A возникла очередь задач B_1, B_2, \dots, B_s , а важно, что в цепочке заданий стоит серия из s экземпляров задачи A . Таким образом, имеем алфавит из имен задач, к которым возможно обращение хотя бы одной информационной технологии. Функционированию нескольких информационных технологий соответствует последовательность задач, решаемых в построенной воображаемой компьютерной системе.

Возникающие очереди в информационной системе соответствуют сериям одного имени задачи в этой последовательности. Необходимость реконфигурации путем установки на резервном хосте дубликата задачи, создающей большую очередь, определяется запретом [6–8], определяемым с помощью порога для длин серий строящейся последовательности. Отсюда возникает следующая вероятностная модель, позволяющая анализировать возможность преодоления защиты информации в сети с помощью задержек в очередях.

Защита от таких нарушений информационной безопасности основана на определении запретов вероятностных мер [10], появление которых инициирует реконфигурацию маршрутов в сети.

Обозначим последовательность задач для решения $x = (x_1, x_2, \dots, x_n, \dots)$, которую строит воображаемая компьютерная система и где $x_i \in \Omega$, $i = 1, 2, \dots$. На множестве Ω определим полиномиальную схему $\{P(A), A \in \Omega\}$. Серия задачи A в последовательности x возникает, когда есть очередь в обращении к задаче A . Серия $C(A, t, l)$ задачи A в последовательности x , начинающаяся с номера t и длиной l , имеет вид:

$$C(A, t, l) = x_t, x_{t+1}, \dots, x_{t+l}, x_{t+l+1},$$

где $x_t \neq A$, $x_{t+1} = \dots = x_{t+l} = A$, $x_{t+l+1} \neq A$. Исходя из модели полиномиальной схемы, вероятность серии $C(A, t, l)$ равна

$$P(C(A, t, l)) = (1 - P(A))^2 P^l(A).$$

Определим случайную величину

$$\xi_t = \begin{cases} 1, & \text{если произошло событие } C(A, t, l); \\ 0 & \text{в противном случае.} \end{cases}$$

Пусть случайная величина T равна времени работы запущенных информационных технологий. Тогда, используя тождество Вальда, можно найти математическое ожидание числа серий

$$\eta(A) = \sum_{t=1}^T \sum_{l=n_0}^{\infty} \xi_t$$

задачи A , длины которых не меньше n_0 :

$$E\eta(A) = (1 - P(A)) P^{n_0}(A) ET.$$

Ясно, что не все задачи могут быть использованы в различных информационных технологиях. Пусть множество этих задач определяется множеством $\Omega_0 \in \Omega$. Тогда среднее число серий задач из Ω_0 , длина которых не меньше n_0 , равно

$$E = \sum_{A \in \Omega_0} E\eta(A) = \sum_{A \in \Omega_0} (1 - P(A)) P^{n_0}(A) ET.$$

Если мощность множества Ω_0 невелика, а $E\eta(A)$, $A \in \Omega_0$, мало, то E является оценкой вероятности появления реконфигурации сети из-за наличия длинной серии запросов хотя бы к одной из задач.

Полученные формулы являются исходными данными для оптимизации задержек одновременного выполнения множества информационных технологий из-за появления очередей при решении последовательности задач.

5 Заключение

В работе построена новая модель, описывающая возможности преодоления защиты сети на основе метаданных при одновременном выполнении нескольких информационных технологий. Рассмотрены две причины задержек:

- (1) задержка в решении задачи одной из информационных технологий;
- (2) появление длинной очереди при одновременном обращении различных технологий к одной и той же задаче.

Такие задержки во многих случаях можно исключать с помощью установки задачи, порождающей очередь, на другой хост (возможно, виртуальный). В этом случае необходима реконфигурация сети с точки зрения построения маршрутов, проходящих через новый хост.

Реконфигурация сети также требует времени, поэтому необходимо оптимизировать размер очереди и момент времени перехода к реконфигурации сети. Построена модель и приведены методы расчета характеристик, связанных с описанным способом преодоления защиты. Полученные формулы позволяют оценивать задержки из-за возникновения очередей в сети или сбоев при решении отдельных задач.

Литература

1. Grusho A., Grusho N., Zabezhailo M., Zatsarinny A., Timonina E. Information security of SDN on the basis of meta data // Computer network security / Eds. J. Rak, J. Bay, I. V. Kotenko, et al. — Lecture notes in computer science ser. — Springer, 2017. Vol. 10446. P. 339–347. https://link.springer.com/chapter/10.1007/978-3-319-65127-9_27.
2. Grusho A. A., Timonina E. E., Shorgin S. Ya. Modelling for ensuring information security of the distributed information systems // 31st European

- Conference on Modelling and Simulation Proceedings. — Dudweiler, Germany: Digitaldruck Pirrot GmbH, 2017. P. 656–660. http://www.scs-europe.net/dlib/2017/ecms2017acceptedpapers/0656-probstat_ECMS2017_0026.pdf.
3. *Grusho A., Timonina E., Shorgin S.* Security models based on stochastic meta data // *Analytical and computational methods in theory probability* / Eds. V. Rykov, N. Singpurwalla, A. Zubkov. — Lecture notes in computer science ser. — Springer, 2017. Vol. 10684. P. 388–400. https://link.springer.com/chapter/10.1007/978-3-319-71504-9_32.
 4. *Грушо А. А., Тимонина Е. Е., Шоргин С. Я.* Иерархический метод порождения метаданных для управления сетевыми соединениями // *Информатика и её применения*, 2018. Т. 12. Вып. 2. С. 44–49.
 5. *Грушо А. А., Забежайло М. И., Зацаринный А. А., Писковский В. О.* Безопасная автоматическая реконфигурация облачных вычислительных сред // *Системы и средства информатики*, 2016. Т. 26. № 3. С. 83–92.
 6. *Grusho A., Timonina E.* Prohibitions in discrete probabilistic statistical problems // *Discrete Math. Appl.*, 2011. Vol. 21. No. 3. P. 275–281.
 7. *Грушо А. А., Грушо Н. А., Тимонина Е. Е.* Статистические методы определения запретов вероятностных мер на дискретных пространствах // *Информатика и её применения*, 2013. Т. 7. Вып. 1. С. 54–57.
 8. *Grusho A., Grusho N., Timonina E.* Power functions of statistical criteria defined by bans // 29th European Conference on Modelling and Simulation Proceedings. — Dudweiler, Germany: Digitaldruck Pirrot GmbH, 2015. P. 617–621. http://www.scs-europe.net/dlib/2015/ecms2015acceptedpapers/0617-probstat_ECMS2015_0016.pdf.
 9. *Gallagher P. R., Jr.* Trusted network interpretation. NCSC-TG-005. Library No. S228.526. Ver. 1 // *The Rainbow Books*, 1987. <https://fas.org/irp/nsa/rainbow/tg005.htm>.
 10. *Grusho A. A., Grusho N. A., Timonina E. E.* Properties of decision functions defined by bans // *J. Math. Sci.*, 2016. Vol. 218. No. 2. P. 154–160.

Поступила в редакцию 02.09.18

OVERCOMING OF PROTECTION OF A NETWORK WHERE CONNECTIONS ARE CONTROLLED BY META DATA

A. A. Grusho, E. E. Timonina, and S. Ya. Shorgin

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

Abstract: The control of connections in a network by means of meta data does not consider time of realization of several information technologies. This fact produces a possibility to overcome protection of a network which uses meta data for control of network connections. For solution of potential problems connected

with this vulnerability, additions to the control functions using meta data are suggested. In particular, the method of creation of queues and the method of compulsory reconfiguration of network are proposed. In the case of SDN (Software Defined Network), meta data directly define the moment of time and requirements to reconfiguration of network for the controller.

Keywords: information security; information technologies; distributed information systems; meta data; software-defined network; composite tasks

DOI: 10.14357/08696527180403

Acknowledgments

The paper was supported by the Russian Science Foundation (project 16-11-10227).

References

1. Grusho, A., N. Grusho, M. Zabezhalo, A. Zatsarinny, and E. Timonina. 2017. Information security of SDN on the basis of meta data. *Computer network security*. Eds. J. Rak, J. Bay, I. V. Kotenko, *et al.* Lecture notes in computer science ser. Springer. 10446:339–347. Available at: https://link.springer.com/chapter/10.1007/978-3-319-65127-9_27 (accessed September 1, 2018).
2. Grusho, A. A., E. E. Timonina, and S. Ya. Shorgin. 2017. Modelling for ensuring information security of the distributed information systems. *31st European Conference on Modelling and Simulation Proceedings*. Dudweiler, Germany: Digitaldruck Pirrot GmbH. 656–660. Available at: http://www.scs-europe.net/dlib/2017/ecms2017acceptedpapers/0656-probstat_ECMS2017_0026.pdf (accessed September 1, 2018).
3. Grusho, A., E. Timonina, and S. Shorgin. 2017. Security models based on stochastic meta data. *Analytical and computational methods in theory probability*. Eds. V. Rykov, N. Singpurwalla, and A. Zubkov. Lecture notes in computer science ser. Springer. 10684:388–400. Available at: https://link.springer.com/chapter/10.1007/978-3-319-71504-9_32 (accessed September 1, 2018).
4. Grusho, A. A., E. E. Timonina, and S. Ya. Shorgin. 2018. Ierarkhicheskiy metod porozhdeniya metadannykh dlya upravleniya setevymi soedineniyami [Hierarchical method of meta data generation for control of network connections]. *Informatika i ee Primeneniya — Inform. Appl.* 12(2):44–49.
5. Grusho, A. A., M. I. Zabezhalo, A. A. Zatsarinny, and V. O. Piskovski. 2016. Bezopasnaya avtomaticheskaya rekonfiguratsiya oblachnykh vychislitel'nykh sred [Secure automatic reconfiguration of cloudy computing]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 26(3):83–92.
6. Grusho, A., and E. Timonina. 2011. Prohibitions in discrete probabilistic statistical problems. *Discrete Math. Appl.* 21(3):275–281.
7. Grusho, A. A., N. A. Grusho, and E. E. Timonina. 2013. Statisticheskie metody opredeleniya zapretov veroyatnostnykh mer na diskretnykh prostranstvakh [Statistical methods of definition of the bans of probability measures on discrete spaces]. *Informatika i ee Primeneniya — Inform. Appl.* 7(1):54–57.

8. Grusho, A., N. Grusho, and E. Timonina. 2015. Power functions of statistical criteria defined by bans. *29th European Conference on Modelling and Simulation Proceedings*. Dudweiler, Germany: Digitaldruck Pirrot GmbH. 617–621. Available at: http://www.scs-europe.net/dlib/2015/ecms2015acceptedpapers/0617-probstat_ECMS2015_0016.pdf (accessed September 1, 2018).
9. Gallagher, P. R., Jr. 1987. Trusted network interpretation. NCSC-TG-005. Library No. S228.526. Ver. 1. Available at: <https://fas.org/irp/nsa/rainbow/tg005.htm> (accessed September 1, 2018).
10. Grusho, A. A., N. A. Grusho, and E. E. Timonina. 2016. Properties of decision functions defined by bans. *J. Math. Sci.* 218(2):154–160.

Received September 2, 2018

Contributors

Grusho Alexander A. (b. 1946) — Doctor of Science in physics and mathematics, professor, principal scientist, Institute of Informatics Problems, Federal Research Center “Computer Sciences and Control” of the Russian Academy of Sciences; 44-2 Vavilov Str., Moscow 119133, Russian Federation; grusho@yandex.ru

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

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

МОДЕЛЬ ГИБРИДНОЙ ИНТЕЛЛЕКТУАЛЬНОЙ МНОГОАГЕНТНОЙ СИСТЕМЫ ГЕТЕРОГЕННОГО МЫШЛЕНИЯ ДЛЯ ИНФОРМАЦИОННОЙ ПОДГОТОВКИ ОПЕРАТИВНЫХ РЕШЕНИЙ В РЕГИОНАЛЬНЫХ ЭЛЕКТРИЧЕСКИХ СЕТЯХ*

А. В. Колесников¹, С. В. Листопад²

Аннотация: Рассматриваются основные модели и алгоритмы поддержки гетерогенного конструирования решений в малых коллективах экспертов. Их применение позволяет фасилитатору или лицу, принимающему решения (ЛПР), организовать коммуникацию и диагностику коллективных эффектов, проблем и корректировку группового поведения. Предлагается формализованная модель гибридной интеллектуальной многоагентной системы гетерогенного мышления (ГИМСГМ), реализующей рассмотренные модели и алгоритмы, для эффективного автоматизированного решения проблем, информационной подготовки и поддержки оперативных решений в региональных электрических сетях.

Ключевые слова: гетерогенное мышление; малый коллектив экспертов; гибридная интеллектуальная многоагентная система

DOI: 10.14357/08696527180404

1 Введение

Оперативно-диспетчерское управление динамической региональной энергосистемой характеризуются разнообразием и значительными объемами необходимой информации, взаимозависимостью принимаемых решений, изменчивостью проблемной среды в реальном времени, влиянием человеческого фактора, в частности неверной интерпретацией данных, слепотой к изменениям и др. Например, при ликвидации технологических нарушений возникают проблемные ситуации [1]: предотвращения развития нарушений, исключение травмирования персонала и повреждения оборудования, не затронутого технологическим нарушением; быстрого восстановления энергоснабжения потребителей и нормальных параметров отпускаемой потребителям электроэнергии с целью снижения убытков сетевой компании; создания надежной послеаварийной схемы. Каждая из

*Работа выполнена при поддержке РФФИ (проект 18-07-00448а).

¹Балтийский федеральный университет им. И. Канта; Калининградский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, avkolesnikov@yandex.ru

²Калининградский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, ser-list-post@yandex.ru

них релевантна модели «неоднородная задача» [2], для их решения необходима интеграция знаний экспертов различных специальностей, согласование нескольких критериев оптимальности и учет множества ограничений. Кроме того, при технологических нарушениях на объектах, входящих в зону ответственности нескольких операторов, важно обеспечить единое понимание ими сложившейся аварийной ситуации и скоординированную работу.

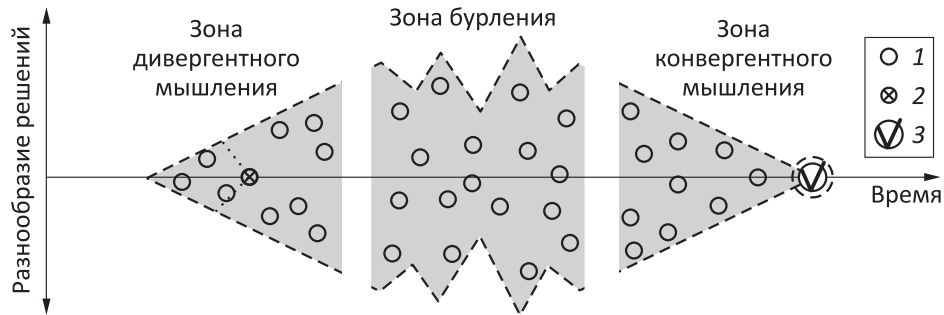
Для моделирования подобных структур в целях информационной подготовки и поддержки принятия решений предлагается подход, интегрирующий гибридный интеллектуальный подход А. В. Колесникова [2], аппарат многоагентных систем в смысле В. Б. Тарасова [3] и методик гетерогенного мышления [4–6]. Результатом должен стать новый класс интеллектуальных систем — ГИМСГМ. Применение ГИМСГМ для информационной подготовки решений позволит автоматизировать деятельность оператора по приему и обработке информации о внешней среде, состоянии системы управления, ходе управляемого процесса, ее анализу, моделированию аварийной ситуации и выработке вариантов ее устранения путем моделирования коллективного решения проблем методами гетерогенного мышления. Результатом такой работы могут выступать новые образы, визуальные формы, несущие определенную смысловую нагрузку [7], позволяющие коллективу диспетчеров увидеть проблему в целом, варианты ее решения, прогноз развития ситуации в каждом случае, принять план действий по устранению аварии и координировать действия при его реализации.

2 Принятие коллективных решений

При решении новых, ранее не встречавшихся задач процесс в общем случае состоит из следующих этапов: постановка, формулирование и анализ задачи, сбор и интерпретация данных, поиск решений, анализ эффективности решений и окончательный выбор, представление результатов, реализация решения, мониторинг и оценка результатов [8]. Процесс решения задачи накладывается на процесс формирования и развития коллектива как единой сущности, состоящий из стадий: формирование, бурление, доработка предложений и подготовка альтернатив, принятие решения и расформирование [9, 10], что согласуется с моделью «ромба группового принятия решений» (см. рисунок) [6].

На первой стадии члены коллектива знакомятся, обмениваются официальной информацией друг о друге, вносят предложения о работе коллектива, придерживаются общепринятых точек зрения, высказывают очевидные решения [10]. Если задача имеет очевидное решение, дискуссия завершается, иначе возникают процессы дивергентного, расходящегося мышления, в рамках которых поощряется безоценочная дискуссия и генерация большого числа решений [6].

Если коллективу удалось выйти за границы устоявшихся мнений, процесс обсуждения переходит в стадию бурления, когда между членами коллектива возможны конфликты из-за противоречивых решений. Под конфликтом будем понимать ситуацию разногласия двух и более экспертов по поводу знаний,



Ромб группового принятия решений С. Кейнера и др. [6]: 1 — альтернатива; 2 — досрочное несогласованное решение; 3 — согласованное решение

убеждений, мнений, т. е. когнитивный конфликт [11]. Конфликт — отличительная черта стадии бурления, позволяющая фасилитатору предпринять меры по выработке взаимопонимания, сближению точек зрения экспертов.

На стадии доработки предложений и подготовки альтернатив члены-эксперты формулируют из ценных мыслей конкретные предложения и «шлифуют» их, пока все участники дискуссии не придут к конечному решению, воплощающему все разнообразие точек зрения. Эта стадия характеризуется «сходящимся мышлением»: классификацией идей, их обобщением, вынесением оценок.

В ходе принятия решений и расформирования вырабатывается коллективное решение, учитывающее мнения всех участников обсуждения.

Модель ромба группового принятия решений может быть использована фасилитатором или его моделью для идентификации сложившейся ситуации принятия решений и попыток направить дискуссию в требуемое русло, активируя соответствующий стиль мышления в коллективе.

3 Формализованная модель гибридной интеллектуальной многоагентной системы гетерогенного мышления

Формально ГИМСГМ определяется следующим образом:

$$\begin{aligned} \text{himsht} &= \langle \text{AG}^*, \text{env}, \text{INT}^*, \text{ORG}, \{\text{ht}\} \rangle; \\ \text{act}_{\text{himsht}} &= \left(\bigcup_{\text{ag} \in \text{AG}^*} \text{act}_{\text{ag}} \right) \cup \text{act}_{\text{dmsa}} \cup \text{act}_{\text{htmc}} \cup \text{act}_{\text{col}}; \\ \text{act}_{\text{ag}} &= (\text{MET}_{\text{ag}}, \text{IT}_{\text{ag}}), \quad \text{ag} \in \text{AG}^*, \quad \left| \bigcup_{\text{ag} \in \text{AG}^*} \text{IT}_{\text{ag}} \right| \geq 2, \end{aligned}$$

где $AG^* = \{ag_1, \dots, ag_n, ag^{dm}, ag^{fc}\}$ — множество агентов, включающее агентов-экспертов (АЭ) ag_i , $i \in \mathbb{N}$, $1 \leq i \leq n$, агента, принимающего решения (АПР), — ag^{dm} и агента-фасилитатора (АФ) — ag^{fc} ; n — число АЭ; env — концептуальная модель внешней среды ГИМСГМ; $INT^* = \{prot, lang, ont, dmscl\}$ — элементы структурирования взаимодействий агентов: $prot$ — протокол взаимодействия; $lang$ — язык передачи сообщений; ont — модель предметной области; $dmscl$ — классификатор ситуаций коллективного решения проблемы, идентифицирующий стадии этого процесса (см. рисунок); ORG — множество архитектур ГИМСГМ; $\{ht\}$ — множество концептуальных моделей макроуровневых процессов в ГИМСГМ: ht — модель процесса коллективного решения проблем методами гетерогенного мышления (см. рисунок); act_{himsht} — функция ГИМСГМ в целом; act_{ag} — функция АЭ из множества AG^* ; act_{dmsa} — функция «анализ ситуации коллективного решения проблемы» АФ; act_{htmc} — функция «выбор метода гетерогенного мышления» АФ; act_{col} — коллективная функция ГИМСГМ, конструируемая динамически; met_{ag} — метод решения задачи; it_{ag} — интеллектуальная технология, в рамках которой реализован метод met_{ag} .

Для реализации функции АФ «анализ ситуации коллективного решения проблемы» act_{dmsa} вводятся понятия совместимости предлагаемых АЭ частных решений, напряженности конфликта между ними и стадии процесса решения проблемы. Совместимость решений — скалярная функция $cmp(dec_k, dec_l) \in [0, 1]$, описывающая возможность одновременной реализации двух частных решений dec_k и dec_l , вид которой зависит от решаемой проблемы.

Напряженность конфликта между двумя агентами может быть определена на основе совместимости предлагаемых ими частных решений проблемы:

$$\begin{aligned} & cnf(ag_i, ag_j) = \\ & = \sum_{k=1}^{N_i} \sum_{l=1}^{N_j} cmp(dec_k, dec_l) (N_i N_j)^{-1}, r_1^{res\ res}(dec_k, ag_i) \circ r_1^{res\ res}(dec_l, ag_j), \end{aligned} \quad (1)$$

где N_i и N_j — количество частных решений, найденных агентами ag_i и ag_j соответственно; $r_1^{res\ res}$ — отношение «быть найденным» между частным решением и предложившим его агентом; \circ — операция склеивания концептов.

Напряженность конфликта в ГИМСГМ в целом описывается выражением:

$$cnf_{himsht} = \sum_{i=1}^n \sum_{j=i+1}^n 2cnf(ag_i, ag_j) (n!)^{-1}. \quad (2)$$

Напряженность конфликта между агентами или в ГИМСГМ используется в качестве универсума лингвистической переменной «конфликт», используемой при реализации функции «выбор метода гетерогенного мышления» act_{htmc} .

Лингвистическая переменная «конфликт» представляется выражением:

$$\text{cnfl} = \langle \beta, T, \text{cnf}, G, M \rangle, \quad (3)$$

где β = «конфликт» — наименование лингвистической переменной; $T = \{ \text{«отсутствует»}; \text{«незначительный»}; \text{«умеренный»}; \text{«острый»} \}$ — терм-множество ее значений, названий нечеткой переменной; $\text{cnf} = [0; 1]$ — универсум нечетких переменных; $G = \emptyset$ — процедура образования из элементов множества T новых термов; $M = \{ \mu_{\text{нет}}(\text{cnf}), \mu_{\text{слабый}}(\text{cnf}), \mu_{\text{умеренный}}(\text{cnf}), \mu_{\text{острый}}(\text{cnf}) \}$ — процедура, ставящая в соответствие каждому терму множества T осмысленное содержание путем формирования нечеткого множества.

Значение символьной переменной «стадия процесса решения проблемы» stg , определенной на множестве $\text{STG} = \{ \text{«дивергентная»}, \text{«бурление»}, \text{«конвергентная»} \}$, вычисляется по правилам:

$$\left. \begin{aligned} \text{stg} = \text{«дивергентная»} \wedge (\text{cnfl} = \text{«умеренный»} \vee \text{cnfl} = \text{«острый»}) &\rightarrow \\ &\rightarrow \text{stg} = \text{«бурление»}; \\ \text{stg} = \text{«бурление»} \wedge (\text{cnfl} = \text{«нет»} \vee \text{cnfl} = \text{«слабый»}) &\rightarrow \\ &\rightarrow \text{stg} = \text{«конвергентная»}. \end{aligned} \right\} \quad (4)$$

Алгоритм функции АФ «анализ ситуаций коллективного решения проблемы» act_{dmsa} — последовательность шагов:

- (1) установить начальные значения: $\text{stg} = \text{«дивергентная»}$, $\text{cnf}(\text{ag}_i, \text{ag}_j) = 0$, $\text{cnf}_{\text{himsht}} = 0$, $\text{cnfl} = \text{«отсутствует»}$;
- (2) ожидать сообщений;
- (3) если получено сообщение о завершении работы ГИМСГМ, то конец;
- (4) если получено сообщение о выработке решения dec_k АЭ ag_i , то переход к п. 5, иначе завершение работы с ошибкой;
- (5) для каждого АЭ ag_j определить $\text{cnf}(\text{ag}_i, \text{ag}_j)$ по формуле (1), где $j \in \mathbb{N}$, $1 \leq j \leq n$, $j \neq i$;
- (6) вычислить $\text{cnf}_{\text{himsht}}$ и cnfl по формулам (2) и (3) соответственно;
- (7) определить стадию процесса решения проблемы stg по правилам (4);
- (8) переход к п. 2.

Функция АФ «выбор метода гетерогенного мышления» act_{htmc} реализуется с использованием нечеткой базы знаний об эффективности методов гетерогенного мышления в зависимости от характеристик проблемы, стадии процесса ее решения и текущей ситуации решения в ГИМСГМ. Эта нечеткая база знаний должна быть разработана по результатам тестирования алгоритмов, реализующих эти методы. Анализ методов гетерогенного мышления [12] позволил для каждой стадии процесса решения проблем (см. рисунок) выделить наиболее перспективные для реализации в ГИМСГМ. Рассмотрим алгоритмы их реализации.

Стадии дивергентного мышления релевантны методы мозгового штурма с наводящими вопросами, пула мозговой записи, оспаривания ограничений и латерального мышления.

Мозговой штурм с наводящими вопросами реализуется следующей последовательностью шагов:

- (1) АПР рассылает АЭ постановку проблемы и, если есть, вариант ее решения;
- (2) АЭ на своих моделях предметной области модифицируют существующий вариант, изменяя способ, место, время и субъект управления;
- (3) найденные решения собираются АПР.

Алгоритм пула мозговой записи:

- (1) АПР рассылает АЭ постановку проблемы;
- (2) каждый АЭ вырабатывает решение и отправляет его АПР;
- (3) АПР, получая решение, создает связанный с ним список АЭ, обработавших его, и записывает в этот список АЭ, нашедшего решение;
- (4) АПР рассылает решения АЭ, которые их не обрабатывали;
- (5) АЭ корректируют решение и отправляют его АПР;
- (6) если каждое решение обработано каждым АЭ, то конец, иначе — п. 4.

Метод оспаривания предположений (ограничений):

- (1) АПР рассылает АЭ постановку проблемы и ее ограничения;
- (2) АЭ вырабатывают решения как с учетом всех ограничений, так и снимая часть из них и отправляют их АПР, указывая нарушенные ограничения;
- (3) АПР собирает решения, чтобы на следующих этапах оценить возможность смягчения ограничений.

Латеральное мышление не алгоритмизируемо по определению, но отдельные механизмы могут быть смоделированы шагами:

- (1) АПР рассылает АЭ постановку проблемы и запрос определенного числа решений (определяется настройками ГИМСГМ) различного качества: лучших и худших по оценке АЭ, а также случайно выбранных;
- (2) АЭ вырабатывают требуемое число решений требуемого качества и отправляют их АПР;
- (3) АПР собирает решения, чтобы на следующих этапах синтезировать из них коллективное решение.

Стадии бурления, согласно [12], релевантны методы взгляда с чужой точки зрения, возврата от решений к потребностям и параллельного мышления.

Взгляд с чужой точки зрения моделируется алгоритмом:

- (1) каждому АЭ случайно назначаются два оппонирующих АЭ;
- (2) оппонирующие АЭ вырабатывают по одному варианту корректировки решений АЭ с указанием причины и отправляют ему;
- (3) АЭ корректирует свои решения на основе своей модели предметной области с учетом рекомендаций;
- (4) АЭ пересылает АПР свои решения.

Метод возврата от решений к потребностям моделируется следующими шагами алгоритма:

- (1) АЭ обмениваются целями, критериями качества решений;
- (2) каждый АЭ модифицирует свой критерий качества с учетом целей других АЭ, при необходимости запрашивает у других АЭ сведения о предметной области для оценки решений в соответствии с новым критерием;
- (3) АЭ вырабатывает решения на своей модифицированной модели предметной области с учетом модифицированного критерия качества;
- (4) АЭ пересылает АПР свои решения.

Моделирование параллельного мышления подразумевает, что АЭ реализуют несколько моделей рассуждений, символизируемых метафорическими шляпами разного цвета: белая — рассуждения на основе информации разной степени надежности; красная — эмоциональное принятие решений; черная — анализ рисков; желтая — «логический позитивизм»; зеленая — дивергентное мышление; синяя — управление мыслительным процессом. Параллельное мышление моделируется алгоритмом:

- (1) АФ просит АЭ «надеть белую шляпу»;
- (2) если АЭ реализует соответствующий шляпе тип рассуждений, он генерирует варианты решений в соответствии с этой моделью рассуждений, иначе он пропускает эту итерацию;
- (3) если все «шляпы» уже надевались, то конец, иначе АФ просит АЭ «надеть» очередную «шляпу», которая ранее не «надевалась», и переход к п. 2.

Стадии конвергентного мышления релевантны поиск консенсуса, голосование и выбор решения ЛПР после обсуждения с членами коллектива [12].

Один из вариантов достижения консенсуса — применение метода Дельфи, который может быть смоделирован следующим алгоритмом:

- (1) АПР рассылает АЭ решения проблемы, полученные на предыдущих стадиях, для оценки;
- (2) АЭ оценивают решения в соответствии со своей моделью предметной области и отправляют оценки АПР;
- (3) определяется медиана и крайние значения оценок;

- (4) АЭ сообщается расположение оценок каждого АЭ, если оценка АЭ сильно отклоняется от медианы, то он должен аргументировать ее или изменить;
- (5) АЭ аргументируют или изменяют свою оценку с объяснением причин корректировки и передают оценку или пояснения АПР;
- (6) определяется коэффициент конкордации Кендалла, если он превышает заданное пороговое значение, то конец, иначе п. 7;
- (7) если число итераций превысило заданное значение, то конец без достижения консенсуса, иначе п. 3.

Алгоритм голосования определяется шагами:

- (1) АПР рассылает АЭ решения проблемы, полученные на предыдущих стадиях, для оценки;
- (2) АЭ оценивают решения в соответствии со своей моделью предметной области и отправляют оценки АПР;
- (3) АПР выбирает решение с максимальной суммарной оценкой.

Алгоритм выбора решения АПР после обсуждения с АЭ:

- (1) АПР рассылает АЭ решения проблемы, полученные на предыдущих стадиях, для оценки;
- (2) каждый АЭ выбирает решение на основе своей модели предметной области, отправляет его АПР вместе с обоснованием;
- (3) АПР при необходимости модифицирует свою модель предметной области на основании обоснований решений АЭ и выбирает решение на основе этой модифицированной модели.

Таким образом, благодаря АФ, иницирующему применение различных методов гетерогенного мышления, и АЭ, реализующим различные технологии искусственного интеллекта, ГИМСГМ динамически перестраивает алгоритм своего функционирования, каждый раз при работе над проблемой вырабатывая релевантный ей гибридный интеллектуальный метод решения. Гибридная интеллектуальная многоагентная система гетерогенного мышления обеспечивает комбинирование репрезентации неоднородной функциональной структуры сложной задачи (проблемы) с гетерогенным коллективным мышлением интеллектуальных агентов, что создает условия для решения проблемы без упрощения и идеализации в динамической среде региональных электросетей.

4 Заключение

В работе рассмотрены основные стадии решения проблем малым коллективом экспертов, выделены релевантные этим стадиям стили мышления участников. Предложены формализованная модель ГИМСГМ и алгоритмы, моделирующие процессы гетерогенного мышления, для релевантного моделирования процесса

решения проблем малым коллективом экспертов. Предложенная ГИМСГМ перемещает имитацию коллективной работы в область синергетической информатики, когда для получения результата, большего, чем сумма работ, выполняемых индивидуально отдельными агентами, необходимо их взаимодействие. Это приводит к самоорганизующимся, социальным моделям управления, каждый элемент которых развивается, получая данные и знания от других элементов, что снижает затраты на разработку и эксплуатацию системы. Применение таких систем в оперативно-диспетчерском управлении позволит вырабатывать решения, релевантные проблемам, которые возникают в динамических средах региональных электросетей.

Литература

1. *Скрябин Н. П., Поздеев Н. Д., Алюнов А. Н.* Оперативное управление распределительными электрическими сетями. — Вологда: ВоГТУ, 2011. 123 с.
2. *Колесников А. В.* Гибридные интеллектуальные системы. Теория и технология разработки. — СПб.: СПбГТУ, 2001. 711 с.
3. *Тарасов В. Б.* От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. — М.: Эдиториал УРСС, 2002. 352 с.
4. *Gardner H.* Multiple intelligences — the theory in practice. — New York, NY, USA: Basic Books, 1993. 320 p.
5. *De Bono E.* Parallel thinking: From Socratic to De Bono thinking. — Melbourne: Penguin Books, 1994. 228 p.
6. *Kaner S., Lind L., Toldi C., Fisk S., Beger D.* The facilitator's guide to participatory decision-making. — San Francisco, CA, USA: Jossey-Bass, 2011. 368 p.
7. *Душков Б. А., Королёв А. В., Смирнов Б. А.* Основы инженерной психологии. — М.: Академический проспект, 2002. 352 с.
8. *Самсонова М. В., Ефимов В. В.* Технология и методы коллективного решения проблем. — Ульяновск: УлГТУ, 2003. 152 с.
9. *Занковский А. Н.* Организационная психология. — 2-е изд. — М.: Флинта, МПСИ, 2002. 648 с.
10. Организационное поведение / Под ред. Г. Р. Латфуллина, О. Н. Громовой. — СПб: Питер, 2004. 432 с.
11. *Tang A. Y. C., Basheer G. S.* A conflict resolution strategy selection method (ConfRSSM) in multi-agent systems // Int. J. Adv. Computer Sci. Appl., 2017. Vol. 8. No. 5. P. 398–404.
12. *Колесников А. В., Листопад С. В., Румовская С. Б., Майтаков Ф. Г.* Анализ методов гетерогенного мышления и перспектив их реализации гибридными интеллектуальными многоагентными системами // Вестник Балтийского федерального университета им. И. Канта. Сер. Физико-математические и технические науки, 2018. № 2. С. 59–71.

Поступила в редакцию 01.09.18

**MODEL OF A HYBRID INTELLIGENT MULTIAGENT SYSTEM
OF HETEROGENEOUS THINKING FOR PREPARATION
OF INFORMATION ABOUT OPERATIONAL DECISIONS
IN A REGIONAL POWER SYSTEM**

A. V. Kolesnikov^{1,2} and S. V. Listopad²

¹Immanuel Kant Baltic Federal University, 14 A. Nevskogo Str., Kaliningrad 236041, Russian Federation

²Kaliningrad 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 main models and algorithms for heterogeneous decision construction support in teams of experts are considered. Their use allows the facilitator or decision maker to organize communication in a team and diagnose collective effects, problems, and adjust team behavior. A formalized model of a hybrid intelligent multiagent system of heterogeneous thinking implementing the considered models and algorithms is proposed for an effective automated problem-solving, information preparation, and support of operational decisions in a regional power system.

Keywords: heterogeneous thinking; expert team; hybrid intelligent multi-agent system

DOI: 10.14357/08696527180404

Acknowledgments

The work was performed with financial support of the Russian Foundation for Basic Research (project 18-07-00448a).

References

1. Skryabin, N. P., N. D. Pozdeev, and A. N. Alyunov. 2011. *Operativnoe upravlenie raspredelitel'nymi elektricheskimi setyami* [Operative management of distribution electric networks]. Vologda: VoGTU. 123 p.
2. Kolesnikov, A. V. 2001. *Gibridnye intellektual'nye sistemy. Teoriya i tekhnologiya razrabotki* [Hybrid intelligent systems: Theory and technology of development]. St. Petersburg: SPbGTU Publ. 711 p.
3. Tarasov, V. B. 2002. *Ot mnogoagentnykh sistem k intellektual'nym organizatsiyam: filosofiya, psikhologiya, informatika* [From multiagent systems to intelligent organizations: Philosophy, psychology, and informatics]. Moscow: Editorial URSS. 352 p.
4. Gardner, H. 1993. *Multiple intelligences — the theory in practice*. New York, NY: Basic Books. 320 p.

5. De Bono, E. 1994. *Parallel thinking: From Socratic to De Bono thinking*. Melbourne: Penguin Books. 228 p.
6. Kaner, S., L. Lind, C. Toldi, S. Fisk, and D. Berger. 2011. *The facilitator's guide to participatory decision-making*. San Francisco, CA: Jossey-Bass. 368 p.
7. Dushkov, B. A., A. V. Korolev, and B. A. Smirnov. 2002. *Osnovy inzhenernoy psikhologii* [Fundamentals of engineering psychology]. Moscow: Akademicheskiy prospect. 352 p.
8. Samsonova, M. V., and V. V. Efimov. 2003. *Tekhnologiya i metody kollektivnogo resheniya problem* [Technology and methods of collective problem solving]. Ul'yanovsk: UIGTU. 152 p.
9. Zankovskiy, A. N. 2002. *Organizatsionnaya psikhologiya* [Organizational psychology]. Moscow: Flinta, MPSI. 648 p.
10. Latfullin, G. R., and O. N. Gromova, eds. 2004. *Organizatsionnoye povedeniye* [Organizational behavior]. St. Petersburg: Piter. 432 p.
11. Tang, A. Y. C., and G. S. Basheer. 2017. A conflict resolution strategy selection method (ConfRSSM) in multi-agent systems. *Int. J. Adv. Computer Sci. Appl.* 8(5):398–404.
12. Kolesnikov, A. V., S. V. Listopad, S. B. Rumovskaya, and F. G. Maitakov. 2018. Analiz metodov geterogennogo myshleniya i perspektiv ikh realizatsii gibridnymi intellektual'nymi mnogoagentnymi sistemami [An analysis of heterogeneous thinking technologies and the prospects for their implementation by hybrid intelligent multiagent systems]. *IKBFU's Vestnik Ser. Physics, Mathematics, and Technology*. 2:59–71.

Received September 1, 2018

Contributors

Kolesnikov Alexander V. (b. 1948) — Doctor of Science in technology, professor, Institute of Physical and Mathematical Sciences and Information Technology, Immanuel Kant Baltic Federal University, 14 A. Nevskogo Str., Kaliningrad 236041, Russian Federation; senior scientist, Kaliningrad Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 5 Gostinaya Str., Kaliningrad 236000, Russian Federation, avkolesnikov@yandex.ru

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

ВИЗУАЛЬНЫЙ ЯЗЫК МОДЕЛИРОВАНИЯ ЭФФЕКТА СИНЕРГИИ В ГИБРИДНЫХ ИНТЕЛЛЕКТУАЛЬНЫХ МНОГОАГЕНТНЫХ СИСТЕМАХ

И. А. Кириков¹, А. В. Колесников², С. В. Листонад³

Аннотация: Решение проблем методами гибридной интеллектуальной многоагентной системы (ГиИМАС) — сложный многоэтапный процесс, в котором периоды индивидуальной работы агентов над частями проблемы сменяются периодами переговоров и обмена мнениями. Если ГиИМАС — черный ящик для пользователя, доверие к результатам ее работы серьезно снижается. Чтобы повысить прозрачность системы и поднять доверие пользователя к ее рекомендациям, предлагается визуальный язык (алфавит, синтаксические и семантические правила) для придания наглядности коллективным решениям и возникающему при этом синергетическому эффекту.

Ключевые слова: малый коллектив экспертов; синергия; гибридная интеллектуальная многоагентная система; визуальный язык

DOI: 10.14357/08696527180405

1 Введение

Обобщение многолетнего опыта разработки и применения ГиИМАС к решению всевозможных проблем [1–3] позволило выявить их существенный недостаток: при взаимодействии с пользователями и специалистами по отладке применяются только символично-логические методы представления информации, практически не активируя визуально-образные, правосторонние рассуждения лица, принимающего решения. Визуализация образа проблемы и процесса выработки ее решений ГиИМАС обеспечит понимание того, как было получено итоговое решение, что повысит доверие к результатам работы ГиИМАС. Визуализация способна запустить механизмы интуитивного, инсайтного принятия решения пользователем, что особенно необходимо при наличии нескольких решений-кандидатов. Кроме того, визуализация сложной структуры системы и узких мест в процессе решения проблемы позволит оперативно вмешиваться в него, производить настройку и отладку системы.

¹Калининградский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, baltbipiran@mail.ru

²Балтийский федеральный университет им. И. Канта, Калининградский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, avkolesnikov@yandex.ru

³Калининградский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, ser-list-post@yandex.ru

Методы визуализации информации для принятия решений обобщены и развиты в работах О. С. Анисимова, Ю. Р. Валькмана, Б. А. Кобринского, О. П. Кузнецова, Г. С. Осипова, Д. А. Поспелова, В. Б. Тарасова, И. Б. Фоминых, Г. П. Щедровицкого, А. Е. Янковской, Р. Арнхейма, Э. Боно, У. Боумена, М. Вертхеймера, Д. Роэма, Д. Сиббета. Для снижения трудоемкости разработки визуальных языков в [4] предложена неформальная аксиоматическая теория ролевых визуальных моделей на основе принципов теории систем и системного анализа. На базе этой теории в данной работе для повышения «прозрачности» ГиИМАС предлагается визуальный язык образного представления решения проблем и возникающего коллективного эффекта — синергии.

2 Понятие визуального языка

В качестве обобщения результатов работ по имитации рассуждений на визуальных образах [5–11] в [4] предложено следующее определение визуального языка как семиотической системы:

$$v1 = \langle VT, VS, VA, VP, v\tau, v\sigma, v\alpha, v\pi \rangle, \quad (1)$$

где VT, VS, VA и VP — множества основных символов, синтаксических, семантических и прагматических правил соответственно; $v\tau$, $v\sigma$, $v\alpha$ и $v\pi$ — правила изменения множеств VT, VS, VA и VP соответственно. Для статического языка, у которого прагматические правила не регламентированы, а определяются разработчиками каждого агента, модель (1) упрощается:

$$v1 = \langle VT, VS, VA \rangle. \quad (2)$$

Множества VT, VS и VA из формулы (2) определяются выражениями:

$$\begin{aligned} VT &= \langle P, D, VR \rangle; \\ VS &= \langle VT, VN, PRU \rangle; \\ VA &= \langle DO, G^{RES}, G^{PR}, G^{ACT}, G^R \rangle, \end{aligned} \quad (3)$$

где

$$\begin{aligned} DO &= \langle RES, PR, ACT, R \rangle; \\ G^{RES} &: RES \rightarrow P; \quad G^{PR} : PR \rightarrow D; \\ G^{ACT} &: ACT \rightarrow P, \quad G^R : R \rightarrow VR. \end{aligned}$$

Здесь P — множество визуальных примитивов; D — множество визуальных измерений, характеризующих визуальные примитивы; VR^n — множество визуальных отношений между одним и более примитивами [7]; VN — словарь нетерминальных символов; PRU — множество продукционных правил; RES, PR и R — множества ресурсов, свойств и отношений соответственно.

3 Алфавит визуального языка моделирования эффекта синергии

Для визуального моделирования синергии в ГиИМАС [12] алфавит языка должен состоять из примитивных форм (см. рис. 1, б–д): сфера, изображающая частное решение проблемы; стрелка, представляющая действие «выработка решения»; кольцо, определяющее «поле» для размещения решений и классифицирующее решения в зависимости от того, каким агентом оно получено и мнения каких агентов использованы для его нахождения; полый цилиндр, высота которого пропорциональна величине эффекта синергии, а цвет указывает на ее знак (зеленый — положительная, красный — отрицательная). Эти примитивы — множества точек (см. рис. 1, а), определяемых визуальными измерениями (см. рис. 2, а–д, в). При визуальном описании примитивов используется визуальное отношение «определение» vr_1^1 (см. рис. 3, а), представляемое прямоугольником, разделенным на две части: в верхней указывается определяемый примитив, а в нижней — определение.

Как показано на рис. 1, а, визуальный примитив «точка» p_1^1 определяется с использованием следующих визуальных измерений: $d_1^1 = \langle d_{1r}^1, d_{1\gamma}^1, d_{ah}^1 \rangle$ — координаты в цилиндрической системе координат (см. рис. 2, а); $d_2^1 = \langle d_{2r}^1, d_{2g}^1, d_{2b}^1 \rangle$ — цвет в цветовой модели RGB (см. рис. 2, б); d_3^1 — прозрачность точки, определяющая часть светового потока, которая проходит через точку, от объектов, расположенных за ней (см. рис. 2, в).

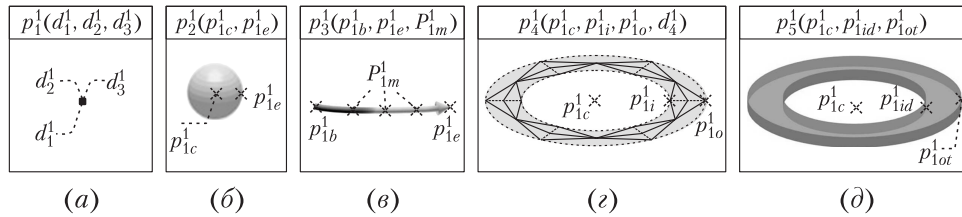


Рис. 1 Множество визуальных примитивов языка моделирования эффекта синергии: (а) точка; (б) сфера; (в) стрелка; (г) кольцо; (д) полый цилиндр

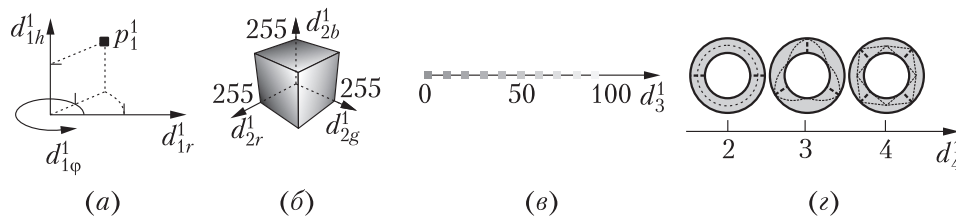


Рис. 2 Множество визуальных измерений языка моделирования эффекта синергии: (а) координаты; (б) цвет; (в) прозрачность; (г) количество агентов

Визуальный примитив «сфера» p_2^1 определяется двумя точками: p_{1c}^1 — центр шара и p_{1e}^1 — точка на поверхности шара. Цвет d_2^1 и прозрачность d_3^1 сферы, т. е. всех ее точек, определяется цветом $d_2^1(p_{1e}^1)$ и прозрачностью $d_3^1(p_{1e}^1)$ точки p_{1e}^1 .

Визуальный примитив «стрелка» p_3^1 определяется не менее чем двумя точками: p_{1b}^1 — начальная узловая точка стрелки и p_{1e}^1 — конечная узловая точка. Кроме того, при наличии непустого кортежа промежуточных узловых точек P_{1m}^1 стрелка представляет собой полилинию. Координаты точек стрелки, лежащих между узловыми точками, определяются методом интерполяции кубическими сплайнами. В случае если цвет и прозрачность двух соседних узловых точек различаются, возникает градиентный переход цвета и прозрачности.

Визуальный примитив «кольцо» p_4^1 однозначно задается точками: p_{1c}^1 — центр кольца, p_{1i}^1 — точка на внутренней окружности кольца и p_{1o}^1 — точка на внешней окружности кольца, а также визуальным измерением d_4^1 — число сегментов, на которое разделено кольцо, и число подсегментов, устанавлива-

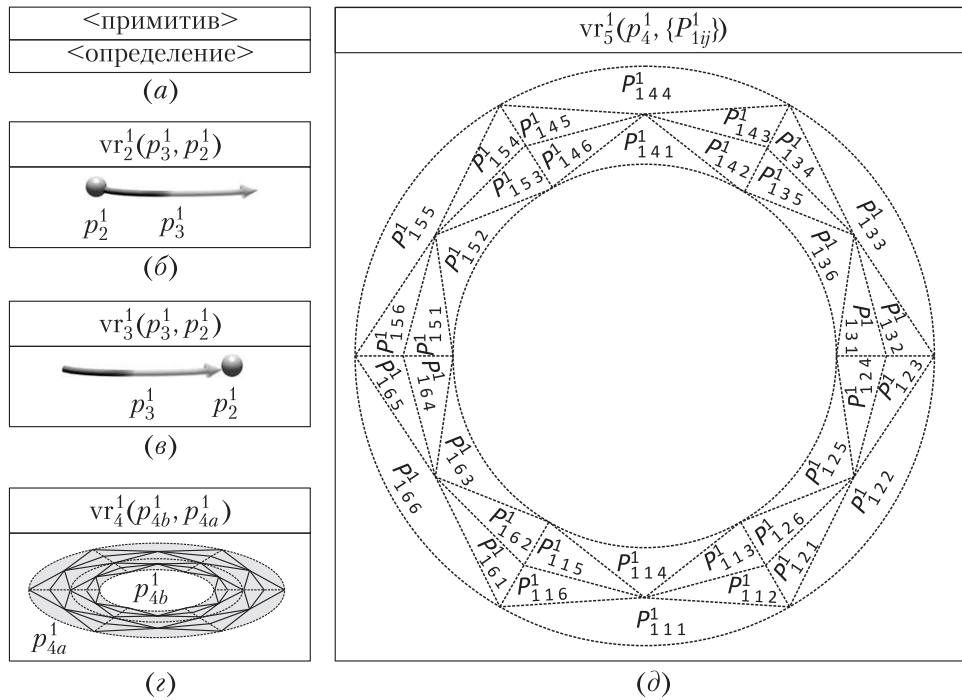


Рис. 3 Множество визуальных отношений языка моделирования эффекта синергии: (а) «определение»; (б) «использовать»; (в) «иметь результатом»; (г) «следовать за»; (д) «иметь частью»

ющих отношения между решением и агентом, получившим его, а также между решением и агентом, предложившим решение — основу для него (см. рис. 2, з). Цвет и прозрачность кольца определяются цветом и прозрачностью точек p_{1i}^1 и p_{1o}^1 . Если их цвет и прозрачность различаются, возникает градиентный переход цвета и прозрачности от центра кольца к краю.

Визуальный примитив «полый цилиндр» p_5^1 однозначно определяется точками: p_{1c}^1 — центр полого цилиндра, p_{1id}^1 — точка на внутренней окружности нижнего кольца полого цилиндра и p_{1ot}^1 — точка на внешней окружности верхнего кольца полого цилиндра. Цвет и прозрачность полого цилиндра определяются цветом и прозрачностью точек p_{1id}^1 и p_{1ot}^1 . Если их цвет и прозрачность различаются, возникает градиентный переход от центра цилиндра к краю.

Визуальные отношения языка для моделирования эффекта синергии представлены на рис. 3.

Отношение vr_2^1 «использовать» указывает, что при выполнении действия «выработка решения», изображенного стрелкой, агент использовал решение, представленное сферой, с которой установлено отношение. Отношение vr_3^1 «иметь результатом» устанавливает связь между действием «выработка решения» и решением — результатом этого действия, представленным сферой, с которой установлено отношение. Отношение vr_4^1 «следовать за» устанавливается между двумя кольцами, определяющими «поля» для размещения решений последовательных шагов итерационного процесса решения проблем [12]. Отношение vr_5^1 «иметь частью» устанавливается между кольцом, определяющим «поле» для размещения решений, и множеством участков этого поля, представляемых множествами P_{1ij}^1 точек. Число участков на кольце составляет $(d_4^1)^2$, а их расположение в случае с шестью агентами, т.е. $d_4^1 = 6$, показано на рис. 3, д.

4 Синтаксис визуального языка моделирования эффекта синергии

Множество VS синтаксических правил PRU определяет механизмы построения нетерминальных символов VN на основе словаря VT терминальных символов (алфавита). Для моделирования эффекта синергии определим правила:

– окраски сфер, представляющих решения:

$$\begin{aligned} \text{pru}_1 &: r_{\text{nbel}}^{1pp} (p_{2i}^1, P_{2s}^1) \rightarrow r_e^{1dd} (d_2^1 [p_{1e}^1 [p_{2i}^1]], \langle 255, 255, 255 \rangle) ; \\ \text{pru}_2 &: r_{1e}^{1dd} (d_{1h}^1 [p_{2i}^1], d_{1h}^1 [p_{2s \max}^1]) \wedge r_{\text{bel}}^{1pp} (p_{2i}^1, P_{2s}^1) \rightarrow \\ &\rightarrow r_e^{1dd} (d_2^1 [p_{1e}^1 [p_{2i}^1]], \langle 255, 0, 0 \rangle) ; \\ \text{pru}_3 &: r_g^{1dd} (d_{1h}^1 [p_{2i}^1], d_{1h}^1 [p_{2s \max}^1]) \wedge r_{\text{bel}}^{1pp} (p_{2i}^1, P_{2s}^1) \rightarrow \\ &\rightarrow r_e^{1dd} (d_2^1 [p_{1e}^1 [p_{2i}^1]], \langle 0, 255, 0 \rangle) , \end{aligned}$$

где r_{nbel}^{1pp} — отношение «не принадлежит» класса «визуальный примитив–визуальный примитив»; r_{bel}^{1pp} — отношение «принадлежит» того же класса; r_e^{1dd} — отношение «быть равным» класса «визуальное измерение – визуальное измерение»; r_{le}^{1dd} — отношение «быть меньше либо равным» того же класса; r_g^{dd} — отношение «быть больше» того же класса; $p_{2s \max}^1 = \arg \max_{p_{2s}^1} (d_{1h}^1 [p_{1c}^1 [p_{2s}^1]])$, $p_{2s}^1 \in P_{2s}^1 = \{p_2^1 | d_{1r}^1 [p_{1c}^1 [p_2^1]] \in d_{1r}^1 [P_{1ii}^1] \wedge d_{1\varphi}^1 [p_{1c}^1 [p_2^1]] \in d_{1\varphi}^1 [P_{1ii}^1]\}$; $[\]$ — операция «проецирование», т. е. $p_{1c}^1 [p_{2i}^1 (p_{1k}^1, p_{1l}^1)] = p_{1k}^1$; $\langle 255, 255, 255 \rangle$, $\langle 255, 0, 0 \rangle$ и $\langle 0, 255, 0 \rangle$ — цветовые константы, соответствующие белому, красному и зеленому цветам;

- окраски стрелок, изображающих действие «выработка решения»:

$$\begin{aligned} \text{pru}_4 : \text{vr}_2^1 (p_3^1, p_2^1) &\rightarrow r_e^{1dd} (d_2^1 [p_{1b}^1 [p_3^1]], d_2^1 [p_{1e}^1 [p_2^1]]) ; \\ \text{pru}_5 : \text{vr}_3^1 (p_3^1, p_2^1) &\rightarrow r_e^{1dd} (d_2^1 [p_{1e}^1 [p_3^1]], d_2^1 [p_{1e}^1 [p_2^1]]) ; \end{aligned}$$

- построения полых цилиндров, отображающих величину эффекта синергии:

$$\begin{aligned} \text{pru}_6 : r_e^{1dd} (d_3^1 [p_{1id}^1 [p_5^1]], 60) &\wedge r_e^{1dd} (d_3^1 [p_{1ot}^1 [p_5^1]], 60) \wedge \\ &\wedge r_e^{1dd} (d_{1r}^1 [d_1^1 [p_{1id}^1 [p_5^1]]], d_{1r}^1 [d_1^1 [p_{1i}^1 [p_4^1]]]) \wedge \\ &\wedge r_e^{1dd} (d_{1\varphi}^1 [d_1^1 [p_{1id}^1 [p_5^1]]], d_{1\varphi}^1 [d_1^1 [p_{1i}^1 [p_4^1]]]) \wedge \\ &\wedge r_e^{1dd} (d_{1r}^1 [d_1^1 [p_{1ot}^1 [p_5^1]]], d_{1r}^1 [d_1^1 [p_{1o}^1 [p_4^1]]]) \wedge \\ &\wedge r_e^{1dd} (d_{1\varphi}^1 [d_1^1 [p_{1ot}^1 [p_5^1]]], d_{1\varphi}^1 [d_1^1 [p_{1o}^1 [p_4^1]]]) \wedge \\ &\wedge r_e^{1dd} (d_{1r}^1 [d_1^1 [p_{1c}^1 [p_5^1]]], d_{1r}^1 [d_1^1 [p_{1c}^1 [p_4^1]]]) \wedge \\ &\wedge r_e^{1dd} (d_{1\varphi}^1 [d_1^1 [p_{1c}^1 [p_5^1]]], d_{1\varphi}^1 [d_1^1 [p_{1c}^1 [p_4^1]]]) ; \end{aligned}$$

$$\begin{aligned} \text{pru}_7 : r_g^1 (d_{1h}^1 [d_1^1 [p_{1c}^1 [p_{2t \max}^1]]], d_{1h}^1 [d_1^1 [p_{1c}^1 [p_{2s \max}^1]]]) &\rightarrow \\ \rightarrow r_e^{1dd} (d_{1h}^1 [d_1^1 [p_{1id}^1 [p_5^1]]], d_{1h}^1 [d_1^1 [p_{1c}^1 [p_{2s \max}^1]]]) &\wedge \\ \wedge r_e^{1dd} (d_2^1 [p_{1id}^1 [p_5^1]], \langle 0, 255, 0 \rangle) &\wedge \\ \wedge r_e^{1dd} (d_{1h}^1 [d_1^1 [p_{1ot}^1 [p_5^1]]], d_{1h}^1 [d_1^1 [p_{1c}^1 [p_{2t \max}^1]]]) &\wedge \\ \wedge r_e^{1dd} (d_2^1 [p_{1ot}^1 [p_5^1]], \langle 0, 255, 0 \rangle) ; & \end{aligned}$$

$$\begin{aligned} \text{pru}_8 : r_{\text{le}}^1 (d_{1h}^1 [d_1^1 [p_{1c}^1 [p_{2t \max}^1]]], d_{1h}^1 [d_1^1 [p_{1c}^1 [p_{2s \max}^1]]]) &\rightarrow \\ \rightarrow r_e^{1dd} (d_{1h}^1 [d_1^1 [p_{1id}^1 [p_5^1]]], d_{1h}^1 [d_1^1 [p_{1c}^1 [p_{2t \max}^1]]]) &\wedge \\ \wedge r_e^{1dd} (d_2^1 [p_{1id}^1 [p_5^1]], \langle 255, 0, 0 \rangle) &\wedge \end{aligned}$$

$$\wedge r_e^{1dd} (d_{1h}^1 [d_1^1 [p_{1ot}^1 [p_5^1]]], d_{1h}^1 [d_1^1 [p_{1c}^1 [p_{2s\max}^1]]]) \wedge \\ \wedge r_e^{1dd} (d_2^1 [p_{1ot}^1 [p_5^1]], \langle 255, 0, 0 \rangle),$$

где $p_{2t\max}^1 = \arg \max_{p_{2t}^1} (d_{1h}^1 [p_{1c}^1 [p_{2t}^1]])$, $p_{2t}^1 \in P_{2t}^1 = P_{2t}^1 \setminus P_{2s}^1$.

Перечисленные правила вместе с алфавитом языка и формируемыми этими правилами нетерминальными символами составляют синтаксис языка, необходимый для построения схематизированных изображений. Для обеспечения взаимосвязи между схематизированными изображениями и формальными символично-логическими рассуждениями ГиИМАС необходимо определить множество семантических правил.

5 Семантические правила визуального языка моделирования эффекта синергии

Согласно выражению (3) множество VA семантических правил отображает модель предметной области DO на множества визуальных примитивов P , измерений D и отношений VR. Для определения этих отношений рассмотрим часть модели предметной области ГиИМАС — концептуальную модель, построенную на основе неформальной аксиоматической теории [1] и описывающую коллективное построение решений проблемы агентами:

$$\text{dmp} = r_1^{\text{res pr}} (\text{res}_{\text{himas}}, \text{pr}_{\text{nag}}) \circ r_1^{\text{res act}} (\text{res}_{\text{exa}}, \text{act}_{\text{dg}}) \circ r_1^{\text{act res}} (\text{act}_{\text{dg}}, \text{res}_{\text{dec}}) \circ \\ \circ r_2^{\text{act res}} (\text{act}_{\text{dg}}, \text{res}_{\text{dec}}) \circ r_1^{\text{res pr}} (\text{res}_{\text{dec}}, \text{pr}_{\text{src}}) \circ \\ \circ r_1^{\text{res act}} (\text{res}_{\text{dma}}, \text{act}_{\text{est}}) \circ r_1^{\text{act res}} (\text{act}_{\text{est}}, \text{res}_{\text{dec}}) \circ \\ \circ r_1^{\text{res pr}} (\text{res}_{\text{dec}}, \text{pr}_{\text{qua}}) \circ r_2^{\text{act pr}} (\text{act}_{\text{est}}, \text{pr}_{\text{qua}}),$$

где $\text{res}_{\text{himas}}$ — ресурс «ГиИМАС»; pr_{nag} — свойство «число агентов»; res_{exa} — ресурс «агент-эксперт»; act_{dg} — действие «выработка решения»; res_{dec} — ресурс «решение»; pr_{src} — свойство «идентификатор агента»; res_{dma} — ресурс «агент, принимающий решения»; act_{est} — действие «оценка решения»; pr_{qua} — свойство «качество решения»; $r_1^{\text{res pr}}$ — отношение «иметь свойство»; $r_1^{\text{res act}}$ — отношение «выполнять действие»; $r_1^{\text{act res}}$ — отношение «иметь объектом»; $r_2^{\text{act res}}$ — отношение «иметь результатом»; $r_2^{\text{act pr}}$ — отношение «изменять свойство»; \circ — операция склеивания концептов [2].

Для визуализации этой части модели предметной области необходимы следующие отображения:

$$g_1^{\text{res}} : \text{res}_{\text{dec}} \rightarrow p_2^1; \quad g_1^{\text{act}} : \text{act}_{\text{dg}} \rightarrow p_3^1; \\ g_1^{\text{pr}} : \text{pr}_{\text{nag}} \rightarrow d_4^1; \quad g_2^{\text{pr}} : \text{pr}_{\text{qua}} \rightarrow d_{1h}^1;$$

$$\begin{aligned}
 g_3^{\text{pr}} &: \exists \text{act}_{\text{dg}} (r_2^{\text{act res}} (\text{act}_{\text{dg}}, \text{res}_{\text{dec}}) \circ r_1^{\text{res pr}} (\text{res}_{\text{dec}}, \text{pr}_{\text{src } i}) \wedge \\
 &\quad \wedge r_1^{\text{act res}} (\text{act}_{\text{dg}}, \text{res}_{\text{dec}}) \circ r_1^{\text{res pr}} (\text{res}_{\text{dec}}, \text{pr}_{\text{scr } j}) \rightarrow \\
 &\quad \rightarrow d_{1r}^1 [d_1^1 [p_{1c}^1 [p_{2i}^1]]] \in d_{1r}^1 [P_{1ij}^1] \wedge d_{1\varphi}^1 [d_1^1 [p_{1c}^1 [p_{2i}^1]]] \in d_{1\varphi}^1 [P_{1ij}^1] ; \\
 g_4^{\text{pr}} &: \neg \exists \text{act}_{\text{dg}} (r_2^{\text{act res}} (\text{act}_{\text{dg}}, \text{res}_{\text{dec}}) \circ r_1^{\text{res pr}} (\text{res}_{\text{dec}}, \text{pr}_{\text{src } i}) \wedge \\
 &\quad \wedge r_1^{\text{act res}} (\text{act}_{\text{dg}}, \text{res}_{\text{dec}}) \circ r_1^{\text{res pr}} (\text{res}_{\text{dec}}, \text{pr}_{\text{scr } j}) \rightarrow \\
 &\quad \rightarrow d_{1r}^1 [d_1^1 [p_{1c}^1 [p_{2i}^1]]] \in d_{1r}^1 [P_{1ii}^1] \wedge d_{1\varphi}^1 [d_1^1 [p_{1c}^1 [p_{2i}^1]]] \in d_{1\varphi}^1 [P_{1ii}^1] ; \\
 g_1^r &: r_1^{\text{act res}} \rightarrow \text{vr}_2^1 ; \quad g_2^r : r_2^{\text{act res}} \rightarrow \text{vr}_3^1 .
 \end{aligned}$$

Эти семантические правила позволяют наглядно отобразить процесс выработки решения в ГиИМАС путем построения схематизированных изображений, аналогичных показанному на рис. 4. По этим изображениям пользователь может наглядно оценивать эффект синергии, возникающий при решении проблем

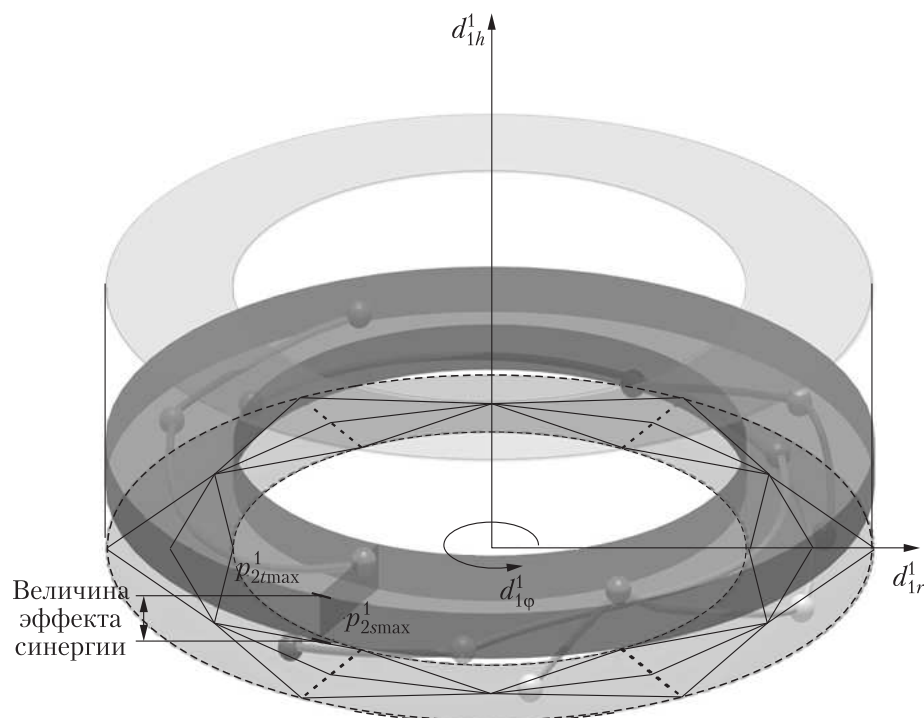


Рис. 4 Пример схематизированного изображения процесса решения проблемы в ГиИМАС и возникающего при этом эффекта синергии

ГиИМАС, по высоте полупрозрачного полого цилиндра и его цвету. Отображение образов-связей между решениями позволяет проследить историю формирования коллективного решения, предлагаемого ГиИМАС, что повышает доверие пользователя к рекомендациям компьютерного искусственного интеллекта.

Схема распределения решений в схематизированных изображениях на рис. 4 позволяет выделить кластеры решений для распознавания и понимания эффективности работы агентов и их групп и выработки релевантных управляющих воздействий на ГиИМАС.

6 Заключение

В работе предложена формальная модель визуального языка как семиотической системы для моделирования эффекта синергии, состоящего из алфавита синтаксических и семантических правил. Предложенный визуальный язык обеспечивает формирование схематизированных изображений, наглядно демонстрирующих возникновение эффекта синергии в ГиИМАС, его знак и величину. Кроме того, благодаря расположению всех элементов визуализации в пределах кольца обеспечивается релевантность метафоре процессов коллективной работы над проблемной ситуацией за круглым столом.

Отображение динамики компьютерной симуляции решения проблемы ГиИМАС, наглядно-образное отображение взаимосвязей предложенных экспертами решений и вклада агентов в итоговое коллективное решение позволяют пользователю ГиИМАС интуитивно принимать решение о необходимости корректировки ее состава или архитектуры, задействуются визуально-образные, правосторонние рассуждения. Визуализация процесса выработки решений ГиИМАС обеспечивает понимание того, как было получено итоговое решение, что повышает доверие к результатам работы системы.

Литература

1. Колесников А. В. Гибридные интеллектуальные системы. Теория и технология разработки. — СПб.: СПбГТУ, 2001. 711 с.
2. Колесников А. В., Кириков И. А., Листопад С. В., Румовская С. Б., Доманицкий А. А. Решение сложных задач коммивояжера методами функциональных гибридных интеллектуальных систем / Под ред. А. В. Колесникова. — М.: ИПИ РАН, 2011. 295 с.
3. Колесников А. В., Кириков И. А., Листопад С. В. Гибридные интеллектуальные системы с самоорганизацией: координация, согласованность, спор. — М.: ИПИ РАН, 2014. 189 с.
4. Колесников А. В., Листопад С. В., Румовская С. Б., Данишевский В. И. Неформальная аксиоматическая теория ролевых визуальных моделей // Информатика и её применения, 2016. Т. 10. Вып. 4. С. 114–120.
5. *Воэтан W. J.* Graphic communication. — New York, NY, USA: John Wiley, 1968. 210 p.

6. *Lakin F.* Visual grammars for visual languages // 6th National Conference on Artificial Intelligence Proceedings. — Menlo Park, CA, USA: AAAI Press, 1987. P. 683–688.
7. *Narayanan N. H., Hubscher R.* Visual language theory: Towards a human–computer interaction perspective // Visual language theory. — New York, NY, USA: Springer-Verlag, 1998. P. 81–128.
8. *Осунюв Г. С.* От ситуационного управления к прикладной семиотике // Новости искусственного интеллекта, 2002. № 6(54). С. 3–7.
9. *Sibbet D.* Visual leaders: New tools for visioning, management, and organization change. — Hoboken, NJ, USA: Wiley, 2013. 229 p.
10. *Fitriani S., Rothkrantz L. J. M.* Two-dimensional visual language grammar. <http://mmi.tudelft.nl/pub/siska/TSD%202DVisLangGrammar.pdf>.
11. *Kremer R.* Visual languages for knowledge representation. <http://ksi.cpsc.ucalgary.ca/KAW/KAW98/kremer>.
12. *Кириков И. А., Колесников А. В., Листопад С. В.* Визуально-образное моделирование эффекта синергии в гибридных интеллектуальных многоагентных системах // Системы и средства информатики, 2018. Т. 28. № 3. С. 165–174.

Поступила в редакцию 01.09.18

VISUAL LANGUAGE FOR MODELING OF SYNERGY EFFECT IN HYBRID INTELLIGENT MULTIAGENT SYSTEMS

I. A. Kirikov¹, A. V. Kolesnikov^{1,2}, and S. V. Listopad¹

¹Kaliningrad Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 5 Gostinaya Str, Kaliningrad 236000, Russian Federation

²Immanuel Kant Baltic Federal University, 14 A. Nevskogo Str., Kaliningrad 236041, Russian Federation

Abstract: Solving problems by a hybrid intelligent multiagent system is a complex multistage process in which periods of agents’ personal work on parts of the problem are followed by periods of negotiations and exchange of opinions. If a hybrid intelligent multiagent system is a black box for the user, the credibility of the results of its work is seriously reduced. To improve the transparency of the system and increase the trust of the user to its recommendations, the authors propose a visual language (alphabet, syntactic and semantic rules) which makes collective decisions and the resulting synergistic effect more visible.

Keywords: expert team; synergy; hybrid intelligent multiagent system; visual language

DOI: 10.14357/08696527180405

References

1. Kolesnikov, A. V. 2001. *Gibridnye intellektual'nye sistemy. Teoriya i tekhnologiya razrabotki* [Hybrid intelligent systems: Theory and technology of development]. St. Petersburg: SPbGTU Publ. 711 p.
2. Kolesnikov, A. V., I. A. Kirikov, S. V. Listopad, S. B. Rumovskaya, and A. A. Domnitskiy. 2011. *Reshenie slozhnykh zadach kommvoyazhera metodami funktsional'nykh gibridnykh intellektual'nykh sistem* [Complex travelling salesman problems solved by the methods of the functional hybrid intelligent systems]. Moscow: IPI RAN. 295 p.
3. Kolesnikov, A. V., I. A. Kirikov, and S. V. Listopad. 2014. *Gibridnye intellektual'nye sistemy s samoorganizatsiey: koordinatsiya, soglasovannost', spor* [Hybrid intelligent systems with self-organization: Coordination, consistency, and dispute]. Moscow: IPI RAN. 189 p.
4. Kolesnikov, A. V., S. V. Listopad, S. B. Rumovskaya, and V. I. Danishevskiy. 2016. *Neformal'naya aksiomaticheskaya teoriya rolevykh vizual'nykh modeley* [Informal axiomatic theory of the role visual models]. *Informatika i ee Primeneniya — Inform. Appl.* 10(4):114–120.
5. Bowman, W. J. 1968. *Graphic communication*. New York, NY: John Wiley. 210 p.
6. Lakin, F. 1987. Visual grammars for visual languages. *6th National Conference on Artificial Intelligence Proceedings*. Menlo Park, CA: AAAI Press. 683–688.
7. Narayanan, N. H., and R. Hubscher. 1998. Visual language theory: Towards a human-computer interaction perspective. *Visual language theory*. New York, NY: Springer-Verlag. 81–128.
8. Osipov, G. S. 2002. *Ot situatsionnogo upravleniya k prikladnoy semiotike* [From situational management to applied semiotics]. *Novosti iskusstvennogo intellekta* [Artificial Intelligence News] 6(54):3–7.
9. Sibbet, D. 2013. *Visual leaders: New tools for visioning, management, and organization change*. Hoboken, NJ: Wiley. 229 p.
10. Fitrianie, S., and L. J. M. Rothkrantz. Two-dimensional visual language grammar. Available at: <http://mmi.tudelft.nl/pub/siska/TSD%20DVisLangGrammar.pdf> (accessed July 15, 2018).
11. Kremer, R. Visual languages for knowledge representation. Available at: <http://ksi.cpsc.ucalgary.ca/KAW/KAW98/kremer/> (accessed July 15, 2018).
12. Kirikov, I. A., A. V. Kolesnikov, and S. V. Listopad. 2018. *Vizual'no-obraznoe modelirovanie efekta sinergii v gibridnykh intellektual'nykh mnogoagentnykh sistemakh* [Visual simulation of synergy effect in hybrid intelligent multiagent systems]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(3):165–174.

Received September 1, 2018

Contributors

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

Kolesnikov Alexander V. (b. 1948) — Doctor of Science in technology; professor, Institute of Physical and Mathematical Sciences and Information Technology, Immanuel Kant Baltic Federal University, 14 A. Nevskogo Str., Kaliningrad 236041, Russian Federation; senior scientist, Kaliningrad Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 5 Gostinaya Str, Kaliningrad 236000, Russian Federation, avkolesnikov@yandex.ru

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

АПРИОРНОЕ ОБРАТНОЕ ГАММА-РАСПРЕДЕЛЕНИЕ В БАЙЕСОВСКИХ МОДЕЛЯХ МАССОВОГО ОБСЛУЖИВАНИЯ*

А. А. Кудрявцев¹, С. И. Палионная², В. С. Шоргин³

Аннотация: Рассматривается байесовский подход к построению модели массового обслуживания $M|M|1|0$. В условиях неопределенности интенсивностей входящего потока и обслуживания исследуются характеристики параметров загрузки больших совокупностей систем обслуживания или одной системы с меняющимися параметрами функционирования. Предполагается, что априорные распределения основных параметров модели известны. Статья продолжает ряд работ авторов, посвященных исследованию байесовских моделей массового обслуживания и надежности. В предположении, что одно из априорных распределений интенсивностей входящего потока и обслуживания является обратным гамма-распределением, а второе — распределением Фреше, вычисляются плотность, функция распределения и моменты коэффициента загрузки. Результаты формулируются в терминах гамма-экспоненциальной функции и могут применяться в различных прикладных задачах для исследования отношений двух независимых случайных величин, имеющих обратное гамма-распределение.

Keywords: байесовский подход; обратное гамма-распределение; распределение Фреше; гамма-экспоненциальная функция; системы массового обслуживания; смешанные распределения

DOI: 10.14357/08696527180406

1 Введение

В рамках теории массового обслуживания получено большое число результатов, важных как с математической, так и с прикладной точки зрения. В настоящее время одним из направлений дальнейшего развития этой теории является рассмотрение более сложных случаев строения систем в целях получения адекватных моделей функционирования сложных агрегатов. В связи с этим исследуются разнообразные вероятностные структуры, основанные в том числе и на рандомизации основных параметров системы.

*Работа выполнена при частичной финансовой поддержке РФФИ (проект 17-07-00577).

¹Московский государственный университет им. М. В. Ломоносова, факультет вычислительной математики и кибернетики, nubigena@mail.ru

²Московский государственный университет им. М. В. Ломоносова, факультет вычислительной математики и кибернетики, sofialionnaya@gmail.com

³Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, VShorgin@ipiran.ru

В данной работе рассматривается байесовский подход [1] к описанию системы массового обслуживания $M|M|1|0$. Предполагается, что интенсивности входящего потока λ и обслуживания μ случайны, что обусловливается стохастичностью среды, в которой происходит функционирование, вариациями на производстве и другими факторами. При этом предполагается, что априорные распределения интенсивностей известны. Результаты работы системы зависят не столько от конкретных значений параметров λ и μ , сколько от их отношения. Поэтому основной задачей при исследовании байесовских систем обслуживания является изучение вероятностных характеристик рандомизированного коэффициента загрузки ρ , который определяется как отношение параметра входящего потока к параметру обслуживания.

Ниже приводятся результаты для вероятностных характеристик коэффициента загрузки ρ модели $M|M|1|0$ в случае, когда в качестве пары априорных распределений параметров системы λ и μ рассматриваются обратное гамма-распределение и распределение Фреше.

2 Основные результаты

Рассмотрим гамма-экспоненциальную функцию [2]:

$$\text{Ge}_{\alpha,\beta}(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} \Gamma(\alpha k + \beta), \quad x \in \mathbb{R}, \quad \alpha \geq 0, \quad \beta > 0.$$

Для дальнейших вычислений понадобится следующее утверждение.

Лемма 1. Пусть $\alpha, \theta, r, s > 0$. Тогда

$$\int_0^{\infty} y^{-r-1} e^{-\alpha/y - (\theta/y)^s} dy = \begin{cases} \theta^{-r} \frac{\text{Ge}_{1/s, r/s}(-\alpha/\theta)}{s}, & s > 1; \\ \alpha^{-r} \text{Ge}_{s, r} \left(-\left(\frac{\theta}{\alpha}\right)^s \right), & s < 1; \\ (\alpha + \theta)^{-r} \Gamma(r), & s = 1. \end{cases}$$

Доказательство. Рассмотрим случай $s > 1$. Используя теорему Лебега о предельном переходе, получаем:

$$\begin{aligned} \int_0^{\infty} y^{-r-1} e^{-\alpha/y - (\theta/y)^s} dy &= \alpha^{-r} \int_0^{\infty} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} t^{r+k-1} e^{-(\theta t/\alpha)^s} dt = \\ &= \frac{\theta^{-r}}{s} \sum_{k=0}^{\infty} \frac{(-\alpha/\theta)^k}{k!} \int_0^{\infty} z^{(r+k)/s-1} e^{-z} dz = \frac{\theta^{-r}}{s} \sum_{k=0}^{\infty} \frac{(-\alpha/\theta)^k}{k!} \Gamma\left(\frac{r+k}{s}\right). \end{aligned}$$

Случай $s < 1$ рассматривается аналогично. Случай $s = 1$ напрямую следует из определения гамма-функции. Лемма доказана.

Рассмотрим случайную величину ξ , имеющую обратное гамма-распределение $IG(q, \theta)$ с плотностью

$$f_{\xi}(x) = \frac{\theta^q e^{-\theta/x}}{\Gamma(q)x^{q+1}}, \quad \theta > 0, \quad q > 0, \quad x > 0,$$

и случайную величину η , имеющую распределение Фреше $Fr(u, \alpha)$ [3] с нулевым минимумом, плотность которого имеет вид:

$$f_{\eta}(x) = \frac{u\alpha^u e^{-(\alpha/x)^u}}{x^{u+1}}, \quad \alpha > 0, \quad u > 0, \quad x > 0.$$

Легко видеть, что для случайных величин ξ и η выполняется следующее утверждение.

Лемма 2. Для случайных величин ξ и η , имеющих обратное гамма-распределение $IG(q, \theta)$ и распределение Фреше $Fr(u, \alpha)$ соответственно, для $z \in \mathbb{R}$ выполняются следующие соотношения:

$$E\xi^z = \frac{\theta^z \Gamma(q-z)}{\Gamma(q)}, \quad z < q; \quad E\eta^z = \alpha^z \Gamma\left(1 - \frac{z}{u}\right), \quad z < u.$$

Леммы 1 и 2 дают возможность вычислить вероятностные характеристики коэффициента загрузки ρ в байесовской модели обслуживания $M|M|1|0$.

Теорема 1. Пусть параметр входящего потока λ имеет обратное гамма-распределение $IG(q, \theta)$, а параметр обслуживания μ имеет распределение Фреше $Fr(u, \alpha)$, причем λ и μ независимы. Тогда при $x > 0$ плотность, функция распределения и моменты коэффициента загрузки $\rho = \lambda/\mu$ имеют вид:

$$f_{\rho}(x) = \begin{cases} \left(\frac{\theta}{\alpha}\right)^q x^{-q-1} \frac{\text{Ge}_{1/u, q/u+1}(-\theta/(\alpha x))}{\Gamma(q)}, & u > 1; \\ u \left(\frac{\alpha}{\theta}\right)^u x^{u-1} \frac{\text{Ge}_{u, u+q}(-(\alpha x/\theta)^u)}{\Gamma(q)}, & u < 1; \end{cases}$$

$$F_{\rho}(x) = \begin{cases} \frac{1}{q\Gamma(q)} \int_{(\theta/(\alpha x))^q}^{\infty} \text{Ge}_{1/u, q/u+1}(-z^{1/q}) dz, & u > 1; \\ 1 - \frac{\text{Ge}_{u, q}(-(\alpha x/\theta)^u)}{\Gamma(q)}, & u < 1; \end{cases}$$

$$E\rho^z = \frac{(\theta/\alpha)^z \Gamma(q-z) \Gamma(z/u + 1)}{\Gamma(q)}, \quad z < q.$$

При $u = 1$ распределение коэффициента загрузки ρ совпадает с распределением Ломакса [4] с параметрами θ/α и q .

Доказательство. Поскольку

$$f_{\rho}(x) = \int_0^{\infty} \frac{u\alpha^u \theta^q e^{-\theta/(xy) - (\alpha/y)^u}}{\Gamma(q)x^{q+1}y^{q+u+1}} dy,$$

вид плотности $f_{\rho}(x)$ вытекает из леммы 1 для всех $u > 0$.

Для функции распределения ρ при $u > 1$ справедливо для $x > 0$:

$$\begin{aligned} F_{\rho}(x) &= \frac{(\theta/\alpha)^q}{\Gamma(q)} \int_0^x y^{-q-1} \text{Ge}_{1/u, q/u+1} \left(-\frac{\theta}{\alpha y} \right) dy = \\ &= \frac{(\theta/\alpha)^q}{\Gamma(q)} \int_0^x y^{-q-1} \sum_{k=0}^{\infty} \frac{(-1)^k y^{-k}}{(\theta/\alpha)^{-k} k!} \Gamma \left(\frac{k}{u} + \frac{q}{u} + 1 \right) dy = \\ &= \frac{1}{q\Gamma(q)} \int_{(\theta/\alpha)^q x^{-q}}^{\infty} \sum_{k=0}^{\infty} \frac{(-z^{1/q})^k}{k!} \Gamma \left(\frac{k}{u} + \frac{q}{u} + 1 \right) dz. \end{aligned}$$

В случае $u < 1$ имеем:

$$\begin{aligned} F_{\rho}(x) &= \frac{u(\alpha/\theta)^u}{\Gamma(q)} \int_0^x y^{u-1} \sum_{k=0}^{\infty} \frac{(-\alpha y/\theta)^k}{k!} \Gamma(ku + u + q) dy = \\ &= \frac{u}{(\theta/\alpha)^u \Gamma(q)} \sum_{k=0}^{\infty} \frac{(-1)^k}{(\theta/\alpha)^{uk} k!} \Gamma(ku + u + q) \int_0^x y^{uk+u-1} dy = \\ &= \frac{x^u}{(\theta/\alpha)^u \Gamma(q)} \sum_{k=0}^{\infty} \frac{(-1)^k x^{uk}}{(\theta/\alpha)^{uk} k!} \frac{\Gamma(ku + u + q)}{k+1} = 1 - \frac{1}{\Gamma(q)} \sum_{l=0}^{\infty} \frac{(-1)^l x^{ul}}{(\theta/\alpha)^{ul} l!} \Gamma(lu + q). \end{aligned}$$

Для нахождения моментов ρ достаточно воспользоваться независимостью случайных величин λ и μ и леммой 2. Теорема доказана.

Рассмотрим симметричный случай априорных распределений.

Теорема 2. Пусть параметр входящего потока λ имеет распределение Фреше $\text{Fr}(v, \theta)$, а параметр обслуживания μ имеет обратное гамма-распределение $\text{IG}(p, \alpha)$, причем λ и μ независимы. Тогда при $x > 0$ плотность, функция распределения и моменты коэффициента загрузки $\rho = \lambda/\mu$ имеют вид:

$$f_{\rho}(x) = \begin{cases} v \left(\frac{\theta}{\alpha}\right)^v x^{-v-1} \frac{\text{Ge}_{v, v+p}(-(\theta/(\alpha x))^v)}{\Gamma(p)}, & v < 1; \\ \left(\frac{\alpha}{\theta}\right)^p x^{p-1} \frac{\text{Ge}_{1/v, p/v+1}(-\alpha x/\theta)}{\Gamma(p)}, & v > 1; \end{cases}$$

$$F_{\rho}(x) = \begin{cases} \frac{1}{\Gamma(p)} \int_0^{\infty} \text{Ge}_{v, v+p}(-z) dz, & v < 1; \\ \frac{(\alpha x/\theta)^p}{v\Gamma(p)} \text{Ge}_{1/v, p/v} \left(-\frac{\alpha x}{\theta}\right), & v > 1; \end{cases}$$

$$E_{\rho}^z = \frac{(\theta/\alpha)^z \Gamma(p+z) \Gamma(1-z/v)}{\Gamma(p)}, \quad z < v.$$

При $v = 1$ распределение коэффициента загрузки ρ совпадает с распределением Дагума [5] с параметрами $(1, \theta/\alpha, p)$.

Доказательство. Аналогично предыдущей теореме для получения выражения для плотности ρ при всех $v > 0$ достаточно воспользоваться леммой 1.

Для функции распределения ρ при $v < 1$ справедливо для $x > 0$:

$$F_{\rho}(x) = \frac{v(\theta/\alpha)^v}{\Gamma(p)} \int_0^x y^{-v-1} \sum_{k=0}^{\infty} \frac{(-(\theta/(\alpha y))^v)^k}{k!} \Gamma(kv + v + p) dy =$$

$$= \frac{1}{\Gamma(p)} \int_{(\theta/\alpha)^v x^{-v}}^{\infty} \sum_{k=0}^{\infty} \frac{(-z)^k}{k!} \Gamma(kv + v + p) dz.$$

В случае $v > 1$ имеем:

$$F_{\rho}(x) = \frac{(\alpha/\theta)^p}{\Gamma(p)} \int_0^x y^{p-1} \sum_{k=0}^{\infty} \frac{(-\alpha y/\theta)^k}{k!} \Gamma\left(\frac{k}{v} + \frac{p}{v} + 1\right) dy =$$

$$= \frac{(\alpha/\theta)^p}{\Gamma(p)} \sum_{k=0}^{\infty} \frac{(-1)^k}{(\theta/\alpha)^k k!} \Gamma\left(\frac{k}{v} + \frac{p}{v} + 1\right) \frac{x^{k+p}}{k+p} =$$

$$= \frac{(\alpha/\theta)^p x^p}{v\Gamma(p)} \sum_{k=0}^{\infty} \frac{(-1)^k x^k}{(\theta/\alpha)^k k!} \Gamma\left(\frac{k}{v} + \frac{p}{v}\right).$$

Соотношение для моментов следует из леммы 2 и независимости случайных величин λ и μ . Теорема доказана.

3 Заключение

Приведенные в работе результаты в очередной раз демонстрируют удобство использования гамма-экспоненциальной функции $Ge_{\alpha, \beta}(x)$ для исследования отношений независимых случайных величин, имеющих распределения гамма-типа, в тех случаях, когда распределение смеси отлично от распределения Ломакса или Дагума.

Литература

1. Кудрявцев А. А., Шоргин С. Я. Байесовские модели в теории массового обслуживания и надежности. — М.: ФИЦ ИУ РАН, 2015. 76 с.
2. Кудрявцев А. А., Титова А. И. Гамма-экспоненциальная функция в байесовских моделях массового обслуживания // Информатика и её применения, 2017. Т. 11. Вып. 4. С. 104–108.
3. Fréchet M. Sur la loi de probabilité de l'écart maximum // Ann. Soc. Polonaise Math. Année 1927. Tome VI. — Krakow, 1928. P. 93–116.
4. Lomax K. S. Business failures: Another example of the analysis of failure data // J. Am. Stat. Assoc., 1954. Vol. 49. No. 268. P. 847–852.
5. Dagum C. A new model of personal income-distribution-specification and estimation // Econ. Appl., 1977. Vol. 30. No. 3. P. 413–437.

Поступила в редакцию 27.08.18

***A PRIORI* INVERSE GAMMA DISTRIBUTION IN BAYESIAN QUEUING MODELS**

A. A. Kudryavtsev¹, S. I. Palionnaia^{1, 2}, and V. S. Shorgin²

¹Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation

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

Abstract: The Bayesian approach to construction of the queuing model $M|M|1|0$ is considered. Under the condition of the arrival and service rates' uncertainty, load parameters' characteristics for large groups of service systems or one system with varying parameters of functioning are studied. It is assumed that the *a priori* distributions of the model's main parameters are known. The paper extends the work devoted to the study of Bayesian queuing and reliability models. Assuming that one of the *a priori* distributions of the arrival and service rates is the inverse-gamma distribution and the other is the Fréchet distribution, the density, distribution function, and moments of the traffic intensity are calculated. The

results are formulated in terms of a gamma-exponential function and can be used to study the relationship between two independent inverse gamma distributed random variables in various applied problems.

Keywords: Bayesian approach; inverse gamma distribution; distribution of Fréchet; gamma-exponential function; queuing systems; mixed distribution

DOI: 10.14357/08696527180406

Acknowledgments

The work was partly supported by the Russian Foundation for Basic Research (project 17-07-00577).

References

1. Kudryavtsev, A. A., and S. Ya. Shorgin. 2015. *Bayesovskie modeli v teorii massovogo obsluzhivaniya i nadezhnosti* [Bayesian models in mass service and reliability theories]. Moscow: FIC IU RAN. 76 p.
2. Kudryavtsev, A. A., and A. I. Titova. 2017. Gamma-eksponentsial'naya funktsiya v bayesovskikh modelyakh massovogo obsluzhivaniya [Gamma-exponential function in Bayesian queuing models]. *Informatika i ee Primeneniya — Inform. Appl.* 11(4):104–108.
3. Fréchet, M. 1927. Sur la loi de probabilité de l'écart maximum. *Ann. Soc. Polonaise Math.* 6:93–116.
4. Lomax, K. S. 1954. Business failures: Another example of the analysis of failure data. *J. Am. Stat. Assoc.* 49(268):847–852.
5. Dagum, C. 1977. A new model of personal income-distribution-specification and estimation. *Econ. Appl.* 30(3):413–437.

Received August 27, 2018

Contributors

Kudryavtsev Alexey A. (b. 1978) — Candidate of Science (PhD) in physics and mathematics, associate professor, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; nubigena@mail.ru

Palionnaia Sofia I. (b. 1995) — student, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; sofiapalionnaya@gmail.com

Shorgin Vsevolod S. (b. 1978) — Candidate of Science (PhD) in technology, 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; VShorgin@ipiran.ru

ИНТЕРОПЕРАБЕЛЬНОСТЬ КАК НАУЧНО-МЕТОДИЧЕСКАЯ И НОРМАТИВНАЯ ОСНОВА БЕСШОВНОЙ ИНТЕГРАЦИИ ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМ

*А. А. Башлыкова¹, А. А. Зацаринный², А. А. Каменщиков³, С. В. Козлов⁴,
А. Я. Олейников⁵, И. И. Чусов⁶*

Аннотация: Проведен анализ известного опыта создания автоматизированных информационно-телекоммуникационных систем (АИТКС) различного назначения на основе принципов интероперабельности в условиях воздействия внешних и внутренних факторов. Показано, что обеспечение интероперабельности на процессной основе представляет собой наиболее экономически эффективный способ бесшовной интеграции функциональных подсистем (ФП). Предложена комплексная модель интероперабельности и угроз. Обоснованы основные меры по решению проблемы интероперабельности.

Ключевые слова: интероперабельность; автоматизированные информационно-телекоммуникационные системы; критическая информационная инфраструктура; стандарты; сложные системы; бесшовная интеграция; процессный подход

DOI: 10.14357/08696527180407

1 Введение

Автоматизированные информационно-телекоммуникационные системы различного назначения предназначены для обработки, представления, отображения и передачи больших объемов разнородной информации (аудио-, видео- и других типов данных) в сложных условиях радиоэлектронной обстановки и интенсивного воздействия различных факторов. Представляя собой организационно-техническое объединение нескольких ФП, реализованных на разнородных программно-аппаратных платформах, АИТКС могут рассматриваться в качестве основы для

¹Московский технологический университет (МИРЭА), bashlykova_a_a_mirea@mail.ru

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

³Институт радиотехники и электроники им. В. А. Котельникова Российской академии наук, prostonau@mail.ru

⁴Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, sv_kozlov@mail.ru

⁵Институт радиотехники и электроники им. В. А. Котельникова Российской академии наук, olein@cplire.ru

⁶Институт радиотехники и электроники им. В. А. Котельникова Российской академии наук, chusov@cplire.ru

создания многофункциональных интегрированных систем, которые, в свою очередь, с учетом основных положений современной нормативной правовой базы [1] относятся к «критической информационной инфраструктуре» (КИИ).

В ближайшей перспективе современные технологии и системотехнические решения по созданию АИТКС должны составить основу интенсивной информатизации органов государственного управления и автоматизации деятельности должностных лиц и специалистов в их составе на современной технологической основе. Расширение разнообразия ФП в составе АИТКС и повышение требований к организации и качеству информационных процессов в интегрированных системах предопределяют важное значение обеспечения интероперабельности всех компонентов АИТКС. На первый взгляд, с этой целью необходимо использовать согласованные наборы современных зарубежных и гармонизированных с ними отечественных стандартов в области информационно-коммуникационных технологий (ИКТ), или, другими словами, профилей, регламентирующих в едином комплексе сбалансированное применение современных технологий и средств для создания АИТКС. Вместе с тем мировой и отечественный опыт свидетельствуют о том, что интероперабельность применительно к АИТКС, используемой в условиях интенсивного информационного противоборства, требует одновременного решения и проблемы информационной безопасности (ИБ) [2], что подразумевает включение в профиль нормативных документов современных стандартов ИБ. В этой связи проблема обеспечения интероперабельности в АИТКС приобретает комплексный характер, и ее решение целесообразно проводить на качественно новом уровне с использованием современных методов процессного подхода, с разработкой комплексной модели интероперабельности, угроз и последующим обоснованием системы мер по решению проблем ее обеспечения на основе совершенствования процессов в жизненном цикле ФП, определяющих условия для создания интегрированной системы.

2 Интероперабельность, основные понятия и подходы к ее анализу и обеспечению

Согласно ГОСТ Р 55062-2012 и ИСО/МЭК/ИЕЕЕ 24765-2010, интероперабельность представляет собой способность двух или более информационных систем или компонентов к обмену информацией и к использованию информации, полученной в результате обмена. В основе ее достижения лежит использование современных профилей. Обеспечение интероперабельности, по существу, является наиболее экономичным беспшовным способом интеграции разнородных ФП. В ГОСТ Р 55062-2012 приведена эталонная модель интероперабельности и в виде блок-схемы общая методика ее обеспечения. Использование ИТ-стандартов обеспечивает только нижний, «технический» уровень интероперабельности. В наибольшей степени интероперабельность может быть достигнута на семантическом и организационном уровнях. С другой стороны, говоря об интероперабельности АИТКС, нельзя не учитывать, что они, во-первых, являются

элементами системы управления, представляющей собой, как отмечено в [3], взаимоувязанную совокупность органов, пунктов и средств управления, и, во-вторых, они сами создаются на основе комплексного применения технических и технологических решений и средств. В этой связи проблема обеспечения интероперабельности в отношении АИТКС должна рассматриваться с учетом условий и факторов, формируемых системой управления, определяющей требования к АИТКС, а также с учетом реализуемости таких требований на основе доступных при создании АИТКС технико-технологических компонентов.

3 Обоснование процессного подхода к обеспечению интероперабельности компонентов в составе автоматизированных информационно-телекоммуникационных систем

Под конкретные задачи системы управления с учетом прогнозируемых направлений их расширения создаются органы управления, в распоряжение которых выделяются необходимые ресурсы для выполнения поставленных задач. Для обеспечения их деятельности создаются пункты управления, оснащаемые средствами управления (автоматизированными системами управления и телекоммуникационными системами). В интересах рационального использования возможностей элементов системы управления целесообразно рассматривать их взаимодействие на процессном уровне. С этой целью необходимо поставить в соответствие каждому из элементов системы управления характерный для него процесс. Разработанные в [4] предложения по структуре процессов в жизненном цикле системы управления представлены в таблице. В настоящее время методы процессного подхода широко представлены и активно применяются к организационным или бизнес-процессам.

С использованием методов процессного подхода производится анализ и оптимизация деятельности по обеспечению качества в сфере разработки и произ-

Классификация процессов по принадлежности к элементам системы управления

Элементы системы управления	Органы управления	Пункты управления	Средства управления
Характер процессов	Организационные (административные или бизнес-процессы)	Организационно-технические процессы	Технико-технологические процессы
Признак классификации	Взаимодействие органов управления, подразделений, должностных лиц, специалистов между собой без применения средств управления	Взаимодействие подразделений, должностных лиц, специалистов между собой с применением средств управления	Взаимодействие средств управления между собой без участия человека-оператора

водства высокотехнологичных систем. В предлагаемой постановке, связанной с расширением области применения процессных методов, задача создания и развития АИТКС в рамках системы управления обладает новизной. Обычно на основе функционального подхода в интересах обеспечения деятельности органов управления проводится выбор необходимых ресурсов на базе имеющихся способов и средств, их комплексование на существующей научно-методической и технологической основе, когда объединяются различные компоненты системы управления в процессе их интеграции, а их единство обеспечивается посредством создания интерфейсов. С точки зрения организационного процесса, обеспечивающего расширение функционала, цель достигается, но с технико-технологической точки зрения сквозной процесс выполнения интегрированных функций реализуется нерациональным образом.

Такой подход при своей простоте, срокам реализации и экономичности обуславливает существенные ограничения в различных ситуациях применения системы управления, поскольку не в полной мере учитывает не только динамику развития угроз и опасностей в сфере ее ответственности, но и не рассматривает на процессном уровне природу развития негативных факторов и их взаимодействие с внутренними процессами в рамках системы управления. В свою очередь, в условиях интенсивной конкуренции в области создания перспективных АИТКС это оказывает сдерживающее влияние.

В этой связи целесообразно рассмотреть обобщенный механизм влияния внешних факторов на состояние системы управления и ее компонентов. К примеру, используя разработанный в [5] подход к описанию взаимосвязи внешних факторов, действующих на объект, запускаемых ими процессов в рамках объекта, приводящих к изменению его технологических свойств, схема их взаимосвязи может иметь вид, представленный на рис. 1.

Конкретный фактор внешних условий для системы управления запускает организационный процесс в виде формирования дополнительных требований к ее объектам либо изменения условий ее функционирования, в результате чего под воздействием фактора изменяются функциональные свойства системы управления на организационно-техническом уровне или на технико-технологическом уровне (см. таблицу). Анализ и реализация этих требований должны учитывать характер их влияния на организационный процесс. Важно отметить, что АИТКС, являясь ключевым элементом системы управления, определяющим



Рис. 1 Общая схема взаимосвязи факторов, процессов и свойств объектов

в значительной мере ее структуру и топологию, имеет главной целью обеспечение технико-технологической поддержки организации и практической реализации организационно-технических и организационных процессов в системе управления. Она должна рассматриваться как объект, находящийся под комплексным воздействием указанных выше факторов. На примере трех видов факторов различной природы (организационного, организационно-технического и технико-технологического характера) иллюстрируется взаимосвязь факторов, запускаемых ими процессов, которые, в свою очередь, проявляются во взаимодействии между собой и приводят к изменению свойств АИТКС, что, в конечном счете, приводит к изменению и свойств системы управления. При этом необходимо отметить их многонаправленное взаимное влияние.

Технико-технологические процессы представляют особый интерес с точки зрения их анализа, оптимизации и обоснования рациональных решений по их интеграции с организационно-техническими процессами в рамках системы управления с учетом достижения синергетического эффекта. В течение длительного периода в разработках средств управления широко применяется принцип функциональной интеграции, реализация которого обычно сопровождается элементами технической избыточности, приводящей в конечном счете к снижению их надежности и экономичности. Анализ подхода к классификации процессов по их принадлежности к элементам системы управления и по их роли в сквозном

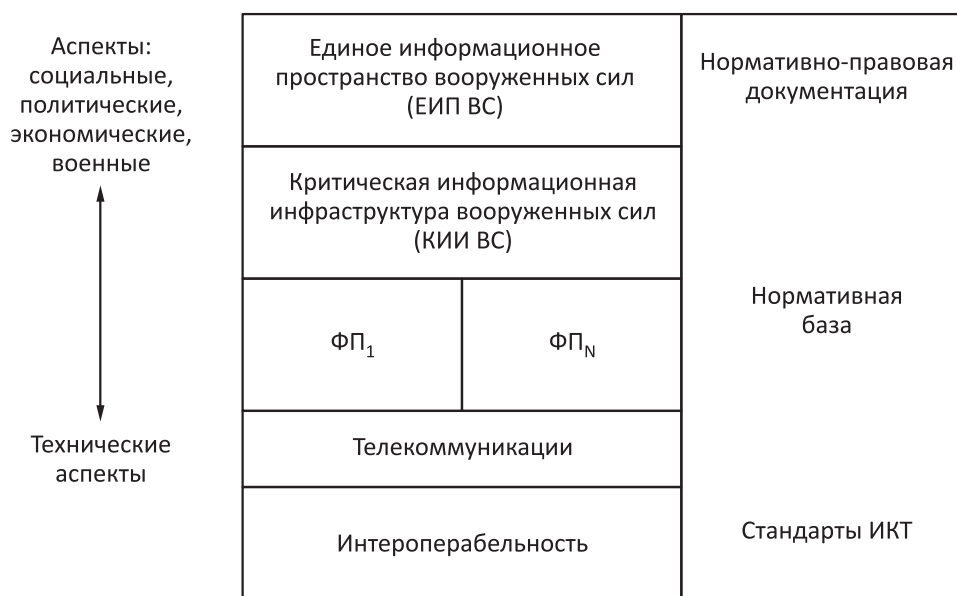


Рис. 2 Соотношение понятий ЕИП, КИИ, интероперабельности и др.

процессе позволяет отметить, что такая декомпозиция процессов создает условия для выявления приоритетности тех или иных процессов при определении основы для интеграции. При этом выбор направления интеграции процессов может осуществляться с учетом направлений взаимосвязи видов процессов по компонентам системы управления:

- от организационных к технико-технологическим через организационно-технические процессы (объединение органов управления и их реорганизация требует изменения средств управления в части расширения их возможностей);
- от технико-технологических процессов через организационно-технические к организационным (появление новых технологий и средств расширяет возможности средств управления, что обуславливает возможность реорганизации органов управления, расширения их функциональных задач).

Совокупность отдельных ФП, объединенных средствами телекоммуникаций, составляет АИТКС, или важнейшую КИИ. На рис. 2 показана взаимосвязь между такими понятиями, как единое информационное пространство (ЕИП) РФ, КИИ, интероперабельность и др. [6].

4 Прикладные аспекты интероперабельности при создании автоматизированной информационно-телекоммуникационной системы

За рубежом на концептуальном и практическом уровне придается важное значение обеспечению интероперабельности при создании АИТКС различного назначения. В нашей стране эта проблема находится в створе актуальных задач развития АИТКС и характеризуется многообразием подходов с использованием существующих протоколов сопряжения. Вместе с тем, несмотря на имеющийся научно-технический и методический задел по ее обеспечению [6, 7], нормативно-методическая база обеспечения интероперабельности в полном объеме не создана. В настоящее время различного рода международные конфликты как с применением способов и средств вооруженного противоборства, так и без него носят характер гибридных войн или, другими словами, войн управляемого хаоса. При этом важным компонентом является информационное противоборство [6]. Несомненно, объектами кибератак служат объекты КИИ, что отмечено в Федеральном законе «О защите критической информационной инфраструктуры». В этих условиях возникает проблема ИБ, отраженная в действующей Доктрине информационной безопасности. Поскольку интероперабельность составляет научно-методическую основу создания АИТКС (см. рис. 2), то объектами атак и объектами защиты должны стать объекты обеспечения интероперабельности. Это означает, что в профиль интероперабельности должны быть включены стандарты ИБ. Для уточнения объектов защиты необходимо на основе модели информационных угроз построить комплексную модель интероперабельности и угроз.

В [6] предложена модель интероперабельности для АИТКС военного назначения, представляющая собой расширение эталонной модели интероперабельности в соответствии с ГОСТ Р 55062-2012 с учетом включения в ее состав такого компонента, как требования ведомственных нормативно-методических документов.

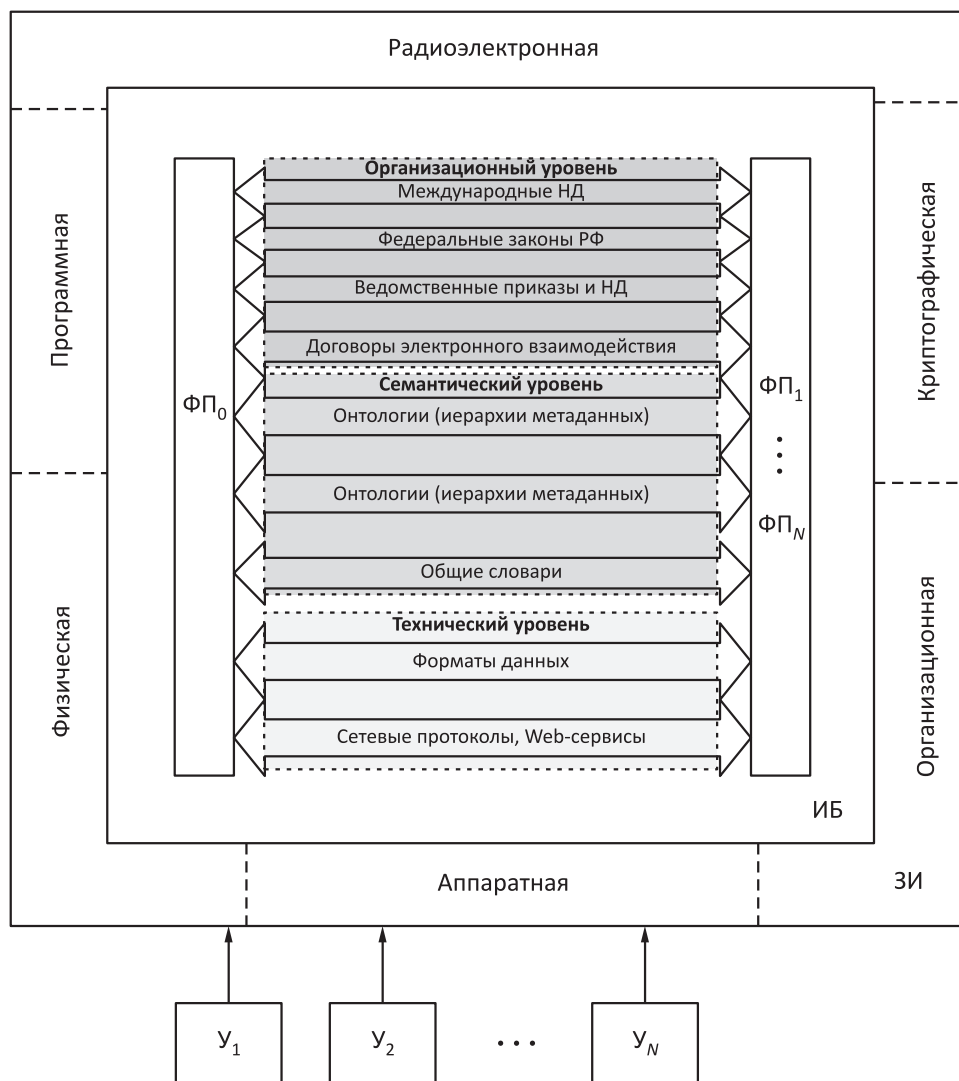


Рис. 3 Комплексная модель интероперабельности и угроз: $ФП_0, ФП_1, \dots, ФП_N$ — функциональные подсистемы; Y_1, Y_2, \dots, Y_N — угрозы ИБ (конфиденциальности, целостности, доступности); ЗИ — виды защиты информации

Затем такая модель с учетом представленной на рис. 1 общей схемы взаимосвязи факторов, процессов и свойств объектов дополняется различного рода факторами, определяющими в конечном счете облик комплексной модели интероперабельности и модели угроз. Важно отметить, что в настоящее время терминология в области интероперабельности и угроз применительно к АИТКС является разнородной. К примеру, в [8, 9] обсуждаются такие понятия, как информационная безопасность, защита информации, информационные угрозы и их классификация, модели информационных угроз, уязвимости, риски, стандарты ИБ. Часто трактовка этих понятий в различных источниках отличается. Предпочтительным представляется использование терминов, содержащихся в стандартах. На основе анализа и обобщения описаний источников угроз ИБ, методов реализации этих угроз, объектов, пригодных для реализации таких угроз, уязвимостей, используемых источниками угроз ИБ, типов возможных потерь (например, нарушение доступности, целостности или конфиденциальности информационных активов), масштабов потенциального ущерба на рис. 3 представлена комплексная модель интероперабельности и угроз.

Достижение интероперабельности в сложных системах представляет собой комплексную проблему, организационно-методические и научно-технические аспекты которой должны разрабатываться одновременно на уровне указанных в таблице процессов на протяжении всех этапов жизненного цикла АИТКС, предусмотренных в стандартах ГОСТ 34.601-90, ГОСТ Р ИСО/МЭК 12207-99 и ГОСТ Р ИСО/МЭК 15288-2005. Следует отметить, что этапы методики обеспечения интероперабельности, приведенные в ГОСТ Р 55062-2012, практически совпадают с этапами и стадиями приведенных стандартов. Таким образом, первая мера состоит в учете процессного подхода при модернизации методики обеспечения интероперабельности на разных этапах жизненного цикла. Вторая мера должна заключаться в создании постоянно действующего рабочего органа по обеспечению интероперабельности, включающего три рабочих группы (технической, семантической и организационной интероперабельности) с представителями заинтересованных ведомств. Пока такой орган не создан, целесообразно использовать возможности подкомитета ПК206/ТК22 «Интероперабельность» [9].

5 Заключение

В соответствии с предложенным подходом к анализу и обеспечению интероперабельности компонентов современных АИТКС:

- достижение интероперабельности в АИТКС представляет комплексную проблему, основные аспекты которой должны разрабатываться одновременно на уровне организационных, организационно-технических и технико-технологических процессов с учетом их развития на всех этапах жизненного цикла АИТКС, с учетом анализа их взаимного влияния и выявления наиболее уязвимых мест на межпроцессном уровне, определяющих возможности обеспечения интероперабельности;

- после декомпозиции процессов с выделением трех групп основных процессов осуществляется переход к комплексной модели интероперабельности и угроз, с помощью которой с использованием основных положений существующих нормативно-методических документов производится обоснование первоочередных мер по обеспечению интероперабельности применительно ко всем этапам жизненного цикла АИТКС;
- разработка профиля стандартов в интересах обеспечения интероперабельности при создании АИТКС должна проводиться одновременно по направлениям трех групп процессов в их жизненном цикле.

Литература

1. О безопасности критической информационной инфраструктуры: Федеральный закон Российской Федерации от 26.07.2017 № 187-ФЗ.
2. *Башлыкова А. А., Олейников А. Я.* Интероперабельность и информационное противоборство в военной сфере // *Ж. радиоэлектроники*, 2016. № 12. 14 с.
3. Основные органы управления, силы и средства ГСЧС. http://texts.news/chrezvychaynyie-situatsii_1537/osnovnyie-organy-i-upravleniya-silyi-sredstva-63396.html.
4. *Козлов С. В.* Об основных направлениях развития телекоммуникационной основы системы управления с учетом комплексного совершенствования организационных, организационно-технических и технико-технологических процессов в ее жизненном цикле // *Радиолокация, навигация, связь: Сб. трудов XXIII Междунар. научно-технич. конф.* — Воронеж: Вэлборн, 2017. Т. 1. С. 147–152.
5. *Топоров А. А., Акусова А. А.* Процессный подход к исследованию изменения технического состояния оборудования химических производств // *Научные труды Донецкого национального технического университета. Сер. Химия и химическая технология*, 2010. Вып. 15(163). С. 153–158.
6. *Каменщиков А. А., Олейников А. Я., Чусов И. И., Ширококова Т. Д.* Проблема интероперабельности в информационных системах военного назначения // *Ж. радиоэлектроники*, 2016. № 11. 16 с.
7. *Зацаринный А. А., Киселев Э. В.* Некоторые подходы к формированию нормативно-технической базы в части требований к архитектурному построению информационных систем организаций участников единого информационного пространства России // *Системы и средства информатики*, 2015. Т. 25. № 3. С. 179–194.
8. *Галатенко В. А.* Стандарты информационной безопасности. — 2-е изд., испр. — М.: Интуит, 2016. 308 с.
9. *Олейников А. Я., Чусов И. И.* Проблема интероперабельности в Вооруженных Силах РФ // *Вестник АВН*, 2017. № 4. С. 61–68.

Поступила в редакцию 10.06.18

**INTEROPERABILITY AS A SCIENTIFIC, METHODOLOGICAL,
AND REGULATORY BASE OF INFORMATION
AND TELECOMMUNICATION SYSTEMS SEAMLESS
INTEGRATION**

*A. A. Bashlykova¹, A. A. Zatsarinny^{2,3}, A. A. Kamenshchikov⁴, S. V. Kozlov³,
A. Ya. Oleynikov⁴, and I. I. Chusov⁴*

¹Moscow Technical University (MIREA), 78 Vernadsky Ave., Moscow 119454, Russian Federation

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

³Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

⁴Kotelnikov Institute of Radioengineering and Electronics (IRE) of the Russian Academy of Sciences, 11-7 Mokhovaya Str., Moscow 125009, Russian Federation

Abstract: The analysis of the known experience in the creation of automated information and telecommunication systems for various purposes based on the principles of interoperability under the influence of external and internal factors is executed. It is shown that providing process-based interoperability is the most cost-effective way of seamless integration of functional subsystems. The authors propose a complex model of interoperability and threats. The main propositions to solve the problem of interoperability are justified.

Keywords: interoperability; automated information and telecommunication systems; critical information infrastructure; standards; complex systems; seamless integration; process approach

DOI: 10.14357/08696527180407

References

1. Federal’nyy zakon Rossiyskoy Federatsii ot 26.07.2017 No. 187-FZ. O bezopasnosti kriticheckoy informatsionnoy infrastruktury [State law of Russian Federation 187-FZ from 26 July, 2017. About critical information infrastructure security].
2. Bashlykova, A. A., and A. Ya. Oleynikov. 2016. Interoperabel’nost’ i informatsionnoe protivoborstvo v voennoy sfere [Interoperability and information confrontation in the military field]. *Zh. radioelektroniki* [Radioelectronics J.] 12. 14 p.
3. Osnovnye organy upravleniya, sily i sredstva GSChS [Key management staff, forces, and means of the SSES]. 2002. Available at: http://texts.news/chrezvyichaynyie-situatsii_1537/osnovnyie-organyi-upravleniya-silyi-sredstva-63396.html (accessed September 11, 2018).

4. Kozlov, S. V. 2017. Ob osnovnykh napravleniyakh razvitiya telekommunikatsionnoy osnovy sistemy upravleniya s uchetom kompleksnogo sovershenstvovaniya organizatsionnykh, organizatsionno-tekhnicheskikh i tekhniko-tekhnologicheskikh protsessov v ee zhiznennom tsikle [About the main directions of development of the telecommunication basis of control system taking into account complex improvement of organizational, organizational and technical and technological processes in its life cycle]. *23rd Scientific and Technical Conference (International) "Radioallocation, Navigation, Communications" Proceedings*. Voronezh: Welborn. 1:147–152.
5. Toporov, A. A., and A. A. Akusova. 2010. Protsessnyy podkhod k issledovaniyu izmeneniya tekhnicheskogo sostoyaniya oborudovaniya khimicheskikh proizvodstv [The process approach to the analysis of changes in the equipment technical state at chemical production]. *Nauchnye trudy Donetskogo natsional'nogo tekhnicheskogo universiteta. Ser. Khimiya i khimicheskaya tekhnologiya* [Scientific Works of Donetsk National Technical University. Ser.: Chemistry and chemical technology] 15(163):153–158.
6. Kamenshchikov, A. A., A. Ya. Oleynikov, I. I. Chusov, and T. D. Shirobokova. 2016. Problema interoperabel'nosti v informatsionnykh sistemakh voennogo naznacheniya [The interoperability problem in information systems of military purposes]. *Zh. radioelektroniki* [Radioelectronics J.] 11. 16 p.
7. Zatsarinnyy, A. A., and E. V. Kiselev. 2015. Nekotorye podkhody k formirovaniyu normativno-tekhnicheskoy bazy v chasti trebovaniy k arkhitekturnomu postroeniyu informatsionnykh sistem organizatsiy uchastnikov edinogo informatsionnogo prostranstva Rossii [Some approaches to the formation of the regulatory and technical base in terms of requirements for the architectural construction of information systems of organization of participators in the unified information space of Russia]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 25(3):179–194.
8. Galatenko, V. A. 2016. *Standarty informatsionnoy bezopasnosti* [Information security standards]. 2nd ed. Moscow: Intuit. 308 p.
9. Oleynikov, A. Ya., and I. I. Chusov. 2017. Problema interoperabel'nosti v Vooruzhennykh Silakh RF [The interoperability problem in the Armed Forces of Russian Federation]. *Vestnik AVN* [Bull. AMS] 4:61–68.

Received June 10, 2018

Contributors

Bashlykova Anna A. (b. 1989) — Candidate of Science (PhD) in technology, associate professor, Moscow Technical University (MIREA), 78 Vernadsky Ave., Moscow 119454, Russian Federation; bashlykova_a_a_mirea@mail.ru

Zatsarinny Alexander A. (b. 1951) — Doctor of Science in technology, professor, Deputy Director, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences (FRC CSC RAS); principal scientist, Institute of Informatics Problems, FRC CSC RAS; 44-2 Vavilov Str., Moscow 119333, Russian Federation; AZatsarinny@ipiran.ru

Kamenshchikov Andrey A. (b. 1982) — Candidate of Science (PhD) in technology, senior scientist, Kotelnikov Institute of Radioengineering and Electronics (IRE) of

the Russian Academy of Sciences, 11-7 Mokhovaya Str., Moscow 125009, Russian Federation; prostonau@mail.ru

Kozlov Sergey V. (b. 1955) — 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 119333, Russian Federation; sv_kozlov@mail.ru

Oleynikov Alexander Ya. (b. 1939) — Doctor of Science in technology, principal scientist, Kotelnikov Institute of Radioengineering and Electronics (IRE) of the Russian Academy of Sciences, 11-7 Mokhovaya Str., Moscow 125009, Russian Federation; olein@cplire.ru

Chusov Igor I. (b. 1945) — Candidate of Science (PhD) in technology, senior scientist, Kotelnikov Institute of Radioengineering and Electronics (IRE) of the Russian Academy of Sciences, 11-7 Mokhovaya St., Moscow 125009, Russian Federation; chusov@cplire.ru

КОНЦЕПТУАЛЬНЫЕ ПОДХОДЫ К СОЗДАНИЮ ИНФОРМАЦИОННО-АНАЛИТИЧЕСКОГО СИТУАЦИОННОГО ЦЕНТРА В ИНТЕРЕСАХ СТРАТЕГИЧЕСКОГО ПЛАНИРОВАНИЯ В ОБЛАСТИ ОБЕСПЕЧЕНИЯ НАЦИОНАЛЬНОЙ БЕЗОПАСНОСТИ

А. П. Сучков¹, А. В. Босов²

Аннотация: Рассмотрены подходы к созданию информационно-аналитического ситуационного центра (ИАСЦ) для управления процессами стратегического планирования в области национальной безопасности. С точки зрения изучения роли ситуационного анализа в этих процессах можно выделить два полных и взаимосвязанных цикла управления, которые условно названы «планирование» и «реализация». Обоснована функциональная модель ИАСЦ, реализующая эти циклы управления, очерчен состав и методы анализа информационных ресурсов, ориентированные на мониторинг, выявление, прогнозирование и оценку угроз. Рассмотрены источники данных, потенциально значимые для подсистемы мониторинга ИАСЦ.

Ключевые слова: стратегическое планирование; ситуационный анализ; цикл управления; информационная модель угрозы; ситуационный центр; информационные источники

DOI: 10.14357/08696527180408

1 Введение

Основным объектом стратегического планирования в области обеспечения национальной безопасности являются угрозы национальной безопасности. Соответствующая деятельность государства составляет предметную область настоящего исследования и представленного далее концептуального подхода. Другие направления стратегического планирования, связанные с вопросами социально-экономического развития, в этой области также важны, но вторичны. В Стратегии национальной безопасности Российской Федерации, утвержденной Указом Президента Российской Федерации от 31 декабря 2015 г. № 683, под угрозой безопасности понимается «совокупность условий и факторов, создающих опасность жизненно важным интересам личности, общества и государства». Государственную деятельность по стратегическому планированию в области обеспечения национальной безопасности осуществляет Совет безопасности РФ. Подход

¹Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, ASuchkov@ipiran.ru

²Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, AVBosov@ipiran.ru

к организации этой деятельности на платформе ситуационного центра (СЦ), которому посвящена данная работа, основывается на представлении предметной области, ее информационной модели и процессах, сформулированных в [1]. Так, ключевым решением является понимание стратегического планирования как циклической последовательности взаимосвязанных процессов, в числе которых два полных и взаимосвязанных цикла управления процессами стратегического планирования, которые будем условно называть «планирование» и «реализация». При этом цикл управления «реализация» явно носит ситуационный характер. В его рамках осуществляется постоянный мониторинг хода выполнения стратегического плана, ситуационный анализ потоков событий контролируемого пространства, текущее реагирование на изменение обстановки вплоть до корректировки документов стратегического планирования.

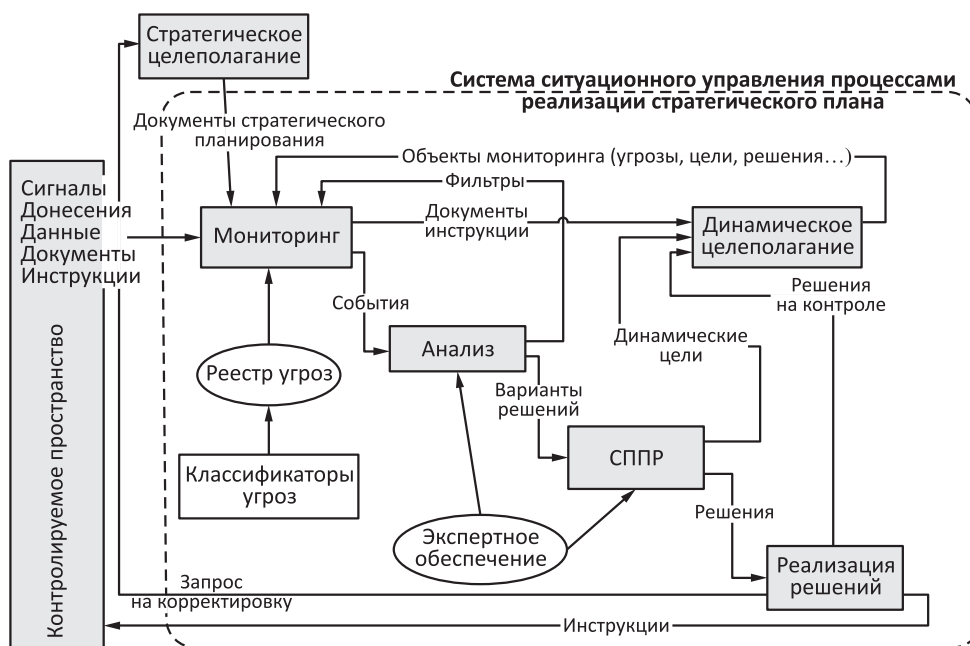
Совбез РФ активно участвует в реализации обоих циклов управления процессами стратегического планирования и в силу этого должен обладать средствами такого управления. Исходя из научно-технического задела и опыта разработки систем ситуационного управления в интересах органов безопасности ФИЦ ИУ РАН, а также руководящих документов высших органов власти в области создания системы распределенных СЦ (СРСЦ), работающих по единому регламенту [2], наиболее адекватной формой реализации процессов управления процессами стратегического планирования является ИАСЦ под эгидой Совбеза.

Ситуационный центр является составной частью системы управления, обеспечивающей информационно-аналитическую поддержку процессов принятия решений в рамках решаемых органом управления функциональных задач на основе автоматизации функций мониторинга и ситуационного анализа обстановки в контролируемом информационном пространстве, включая сбор, обобщение, аналитическую обработку, хранение, передачу, визуализацию и защиту информации.

2 Функциональная модель информационно-аналитического ситуационного центра

Представленная далее функциональная модель ИАСЦ исходит из того, что последовательность процессов, составляющих стратегическое планирование как в рамках цикла «планирование», так и цикла «реализация», включает целеполагание, мониторинг, анализ, принятие решений и их реализацию.

Главным основанием, обосновывающим целесообразность функционирования ИАСЦ в рассматриваемой области, представляется факт наличия процессов *стратегического целеполагания*, т. е. определения направлений, целей и приоритетов обеспечения национальной безопасности Российской Федерации, и его реализация в виде документов стратегического планирования. Система целеполагания ИАСЦ включает в себя также процессы динамического задания целей текущего управления ходом реализации стратегического плана (см. рисунок). Это первый, исходный функциональный блок ИАСЦ.



Функциональная модель системы ситуационного управления

Второй функциональный блок ИАСЦ составляет *система мониторинга*. Базовый перечень объектов мониторинга определен в [1] и включает такие объекты, как угрозы национальной безопасности, целевые показатели, объекты угроз, средства, способы и неконтролируемые факторы их реализации.

Мониторинг угроз опирается на реестр угроз и соответствующие классификаторы, на источники данных мониторинга, включая документы стратегического планирования, на критерии и правила отбора данных из различных информационных ресурсов. В рамках процессов мониторинга осуществляется оценка параметров ненаблюдаемых (скрытых) элементов обстановки на основе выборочных или косвенных данных, анализ неструктурированных данных, выявление и извлечение фактов.

Процессы мониторинга фиксируют события, представляющие собой изменения состояния объектов мониторинга, которые поступают на вход функционального блока *анализа хода реализации документов стратегического планирования*. Здесь осуществляется оценка сложившейся ситуации (штатная, нештатная) и формирование гипотез о состоянии угроз, а также анализ и учет существующего опыта. Процессы анализа опираются на методы декомпозиции складывающихся ситуаций на известные или достаточно простые ситуации для нормализации с помощью ресурсов системы. Наконец, в зависимости от того, какая ситуация

идентифицирована, с использованием экспертного обеспечения вырабатываются варианты реакции ИАСЦ.

Функциональный блок *системы поддержки принятия решений* (СППР) предназначен для формирования решения на основе возможных альтернатив с использованием методов многокритериального выбора. Оценка качества выбранных альтернатив решений может также осуществляться с использованием методов сценарного прогнозирования, что позволяет учесть влияние управляющих воздействий ИАСЦ на развитие обстановки. Блок СППР должен обеспечить также формирование динамических целей для обеспечения процессов мониторинга хода реализации решений.

Процессы блока *реализации решений* обеспечиваются инфраструктурой формализованного электронного документооборота для их доведения до исполнителей и контроля исполнения. Данный функциональный блок при необходимости осуществляет обратную связь с системой стратегического целеполагания с целью корректировки документов стратегического планирования.

Приведенная функциональная модель позволяет сформулировать достаточные функциональные требования и требования по назначению при подготовке технического задания на ИАСЦ.

3 Источники данных мониторинга

Следующий и, возможно, ключевой вопрос функционирования системы стратегического планирования ИАСЦ — это вопрос состава и объемов информации, используемой для мониторинга, и, возможно даже в большей степени, ее источников. Состав источников данных должен определяться функциями, реализуемыми Совбезом РФ при осуществлении основных процессов стратегического планирования в области обеспечения национальной безопасности. К этим функциям относятся:

- выявление внутренних и внешних стратегических рисков, влияющих на устойчивое развитие Российской Федерации, а также угроз национальной безопасности;
- организация мониторинга и проведение обобщенных оценок социально-экономического развития Российской Федерации и состояния национальной безопасности;
- мониторинг показателей состояния национальной безопасности;
- мониторинг эффективности работы по реализации документов стратегического планирования;
- координация деятельности по разработке в федеральных округах документов стратегического планирования.

В Стратегии национальной безопасности Российской Федерации определено, что информационную основу ее реализации «составляет федеральная инфор-

мационная система стратегического планирования, включающая в себя информационные ресурсы органов государственной власти и органов местного самоуправления, системы распределенных ситуационных центров и государственных научных организаций». Основываясь на данном тезисе, а также на опыте создания близких по тематике информационно-аналитических систем (ИАС), выделим следующие группы источников данных для мониторинга ИАСЦ:

- информационные системы федеральных и региональных органов исполнительной власти;
- зарубежные и отечественные базы данных (БД) научно-технической направленности;
- данные новостных систем массовой информации (СМИ);
- информационные ресурсы социального Интернета (социальные сети, мессенджеры, блоги и т. п.),

и рассмотрим наиболее характерных представителей этих групп.

4 Система распределенных ситуационных центров

В соответствии с Концепцией СРСЦ [3] основными целями создания СРСЦ являются: обеспечение информационно-аналитической поддержки государственного управления, *стратегического планирования* и *мониторинга реализации документов стратегического планирования* в Российской Федерации; повышение эффективности государственного управления в мирное и военное время, в том числе при возникновении чрезвычайных (кризисных) ситуаций, за счет использования информационных и технологических возможностей СЦ, обеспечивающих анализ, оценку, прогнозирование изменения обстановки и поддержку принятия управленческих решений.

В рамках создания СРСЦ осуществляется решение следующих задач:

- интеграция действующих СЦ в систему на основе единого регламента;
- создание новых и модернизация действующих СЦ на основе типовых комплексов программно-технических средств с использованием защищенной телекоммуникационной сети;
- формирование распределенного информационного фонда системы, доступ к ресурсам которого должен осуществляться на основе единого рубрикатора и с использованием технологии информационных порталов;
- разработка и внедрение информационных комплексов системы, организация их взаимодействия с информационными комплексами федеральных органов исполнительной власти, иных государственных органов и организаций;
- разработка и внедрение единого регламента взаимодействия, организация управления и обеспечение координации деятельности СЦ на основе этого регламента;

- обеспечение необходимого уровня информационной безопасности информационно-коммуникационной инфраструктуры системы, предусматривающего возможность обмена информацией, содержащей сведения, составляющие государственную тайну, и информацией, в том числе служебной, ограниченного доступа;
- разработка нормативных актов и иных документов, определяющих порядок создания, функционирования и развития системы;
- организация подготовки кадров, включая специалистов в области стратегического планирования, для обеспечения функционирования системы.

Следует отметить, что если процессы интеграции находятся еще в зачаточном состоянии, то создание ведомственных, федеральных окружных и региональных ЦС идет полным ходом на достаточно высоком научно-техническом уровне. Поэтому данная группа источников является одной из основных для обеспечения процессов мониторинга угроз национальной безопасности, контроля реализации документов стратегического планирования и координации деятельности региональных органов государственной власти.

Система межведомственного электронного взаимодействия

Система межведомственного электронного взаимодействия (СМЭВ) представляет собой федеральную государственную информационную систему, включающую информационные базы данных, в том числе содержащие сведения об используемых органами и организациями информационных, программных и технических средствах, обеспечивающих возможность доступа через систему взаимодействия к их информационным системам [4].

Система межведомственного электронного взаимодействия состоит из сети защищенных каналов связи между узлами, расположенными в центрах обработки данных «Ростелекома». Каждый узел СМЭВ — это шина на базе Oracle Enterprise Service Bus.

В 2009 г. обеспечена возможность взаимодействия между Единым порталом государственных услуг [5] и информационными системами федеральных органов исполнительной власти для заказа государственных услуг в электронном виде. Взаимодействие обеспечивается с использованием синхронных электронных сервисов. Третья версия СМЭВ официально введена в эксплуатацию с начала 2015 г.

Задачами СМЭВ в инфраструктуре Электронного правительства РФ являются:

- обеспечение информационного взаимодействия в электронной форме при предоставлении государственных и муниципальных услуг и исполнении государственных и муниципальных функций;
- обеспечение предоставления государственных и муниципальных услуг в электронной форме, в том числе с использованием федеральной государственной

информационной системы «Единый портал государственных и муниципальных услуг (функций)»;

- технологическое обеспечение информационного взаимодействия с применением системы взаимодействия (достигается путем использования сервис-ориентированной архитектуры, представляющей собой совокупность видов сведений, построенных по общепринятым стандартам, а также путем использования единых технологических решений и стандартов, единых классификаторов и описаний структур данных).

Следует отметить, что интеграционное ядро СМЭВ опирается на единые стандарты, классификаторы, описания структур данных, представленные в виде XML-моделей, и охватывает часть данных отраслей народного хозяйства — финансово-экономические, социальные и личные данные.

Нормативно-правовое обеспечение СМЭВ, таким образом, делает его важным источником данных в интересах мониторинга угроз национальной безопасности.

Научно-техническая информация

Важная задача системы мониторинга угроз — выявление их на *стадии замысла и подготовки*, поэтому существенными источниками данных для системы мониторинга угроз национальной безопасности на стадиях замысла и подготовки должны служить отечественные и зарубежные базы данных по передовым научным исследованиям (научным публикациям), по производственной и торговой деятельности зарубежных научно-производственных фирм, а также по изучению опыта действий государственных и частных организаций по предотвращению и ликвидации последствий реализации угроз.

Исходя из сформулированных целей системы выявления, мониторинга и предупреждения угроз, можно выделить следующие необходимые для этого источники данных, представляющие собой зарубежные БД научно-технической информации:

- публикации на английском языке ресурса arXiv.org (1 млн 250 тыс. документов) по естественным наукам;
- англоязычная Википедия (5 млн 164 тыс. статей);
- БД по американским патентам (USPTO с 2002 по 2016 гг.): 2 млн 965 тыс. патентов;
- БД по международным патентам (WIPO): 2 млн 384 тыс. патентов.

Другим значимым источником в этой области являются отечественные и зарубежные электронные реферативные БД:

- БД ВИНТИ РАН;
- электронный каталог ГПНТБ России;
- Scopus (SciVerse Scopus) — библиографический индекс, Elsevier;

- Web of Science — библиографический индекс, Thomson Reuters;
- NTIS (National Technical Information Service) — политематическая БД, U.S. Department of Commerce;
- Life Sciences Collection — БД по естественным наукам, Cambridge Scientific Abstracts;
- Biological Abstracts — БД по биологии, Biological Abstracts Inc.;
- Biotechnology & Bioengineering — БД по биотехнологии и биоинженерии, Cambridge Scientific Abstracts and Engineering Information Inc.;
- Chemical Engineering and Biotechnology Abstracts — БД химической инженерии и биотехнологии, The Royal Society of Chemistry;
- Compendex Plus — БД по техническим наукам, Engineering Information Inc.;
- Corporate & Industry Research Reports (CIRR) — БД аналитических отчетов фирм, JA Micropublishing;
- Current Contents — еженедельное обозрение содержания научных периодических изданий, Institute for Scientific Information;
- Derwent Biotechnology Abstracts — БД по биотехнологии, Derwent Publications Ltd.;
- MEDLINE — БД по медицине и биологии, National Library of Medicine;
- TOXLINE — БД по токсикологии, National Library of Medicine, Swedish Nat. Chemicals Inspectorate.

Не претендуя на полноту и окончательность формирования перечня ресурсов в этом направлении, отметим их тематическое разнообразие, значительные объемы и необходимость наличия глубоких научных компетенций для эффективного использования таких данных.

Новостные системы массовой информации

Данные новостных СМИ лежат в основе мониторинга следующих аспектов противодействия угрозам национальной безопасности:

- оперативное выявление изменений в нормативно-правовой области стратегического планирования в сфере безопасности;
- реакции и отзывы в связи с высказываниями и поручениями высокопоставленных лиц государства в адрес субъектов стратегического планирования;
- освещение деятельности силовых структур и органов правопорядка и безопасности;
- публикации по конкретным инцидентам реализации угроз;
- выявление признаков деятельности, подрывающих национальную безопасность.

Состав новостных СМИ и его изменения характеризуются значительной динамикой в основном только за счет региональных и специализированных изданий (число центральных изданий и их активность в настоящее время почти не меняется). Восемьдесят процентов всех новостей из группы центральных СМИ производят сайты газет, информагентств, телеканалов и радиостанций, Интернет для этих типов СМИ не является единственным каналом вещания. В региональной группе на такие СМИ приходится только половина новостей, другую половину медиапотока создают сетевые СМИ и тематические сайты.

Основные группы новостных СМИ для системы мониторинга угроз:

- сайты крупных информагентств и телеканалов, например ВГТРК, РИА Новости, ИТАР-ТАСС, НТВ, ТВ-Центр, а также множество преимущественно новостных интернет-изданий, среди которых Lifenews.ru, KM.ru, Infox.ru, Ридус.ру и др.;
- региональные и городские СМИ: Fontanka.ru и др.;
- интернет-СМИ: Газета.Ru, Slon.ru, NEWSru.com, Snob.ru, «Эхо Москвы», «Новая газета», телеканал «Дождь» и др.;
- телевизионные каналы и набор разнообразных специализированных СМИ;
- деловые СМИ: «Ведомости», «Коммерсантъ», «РБК» и др., а также ряд специальных изданий, например Право.ру, РИА Новости-Недвижимость, Zoom.cnews.ru, Турпром;
- популярные ежедневные газеты: «Комсомольская правда», «Московский комсомолец», «Аргументы и факты» и др.;
- сайты экстремистских организаций (по соответствующему списку Минюста).

Социальный Интернет

Анализ ресурсов Интернета в целом и его социальной части в частности играет важнейшую роль в мониторинге угроз. Так, агентство DARPA (Defense Advanced Research Project Agency), отвечающее за разработку новых технологий для использования в вооруженных силах, еще в 2001 г. запустило проект «Знания информации о терроризме» (Terrorism Information Awareness — TIA) [6]. Он заключается в создании и испытании опытного образца системы, позволяющей на основе больших объемов несвязанной информации в различных БД (в том числе и в соцсетях) выявить группу лиц, готовящихся совершить террористический акт на территории США. Проект успешно реализован и развивается, а в настоящее время стал фундаментом для последующих разработок в области анализа данных из открытых источников (в том числе и из Интернета).

Другим важным аспектом мониторинга социального Интернета представляется выявление угроз национальной безопасности в информационной сфере, а также осуществление оборонительных и наступательных действий в развязанной в последнее время информационной войне против России. Так, в США

уже несколько лет действуют технологии сквозного мониторинга трафика электронной почты и телефонных звонков, при этом используются методы анализа социальных сетей. Работы идут в рамках проекта Агентства национальной безопасности США под названием Stellar Wind, а также силами контртеррористического подразделения ФБР и других правоохранительных органов в рамках системы Threat and Local Observation Database (TALON).

Приведем примеры, не претендуя на полноту охвата даже ключевых тематик, основных групп ресурсов социального Интернета:

- глобальные мировые социальные сети, блогосферы, СМИ: Google+, Facebook, MySpace, российские: xbeee.com, Placeword.com, ВКонтакте.ру, Одноклассники.ру;
- социальная коммерция: Voiceland;
- деловые контакты: Webby.ru, Профессионалы.ru, LinkedIn, Мой Круг, Aloha Connections;
- менеджмент: E-executive.ru;
- финансы: vProfite.com, Common.ru;
- блог-сервисы, микроблоги: Fixfeel — world emotional map, Twitter, Tip-Topic.net, Рутвит, В основе, Rec Name, LiveJournal — Живой Журнал, Diary.ru;
- общественная жизнь, движения: Vpiski.Net, Нефору.ру, družhno.com, Будист.ру, Mebedik, Страна благодетелей, Politiko, В Политике, Politix.ru, MyChurch, Моя Православная Сеть.

5 Аналитика информационно-аналитического ситуационного центра и дальнейшие перспективы

Информационные и функциональные модели, тематика и состав источников информационного обеспечения — достаточные основания для проработки следующего уровня решений: технических и технологических. Этой составляющей ИАСЦ здесь коснемся кратко, затронув только самый значимый компонент, точнее самую насущную информационную технологию ИАСЦ — ИАС. Информационно-аналитическая система в составе ИАСЦ обеспечит ключевую технологическую платформу обработки данных мониторинга, даст средства автоматизации деятельности экспертно-аналитического и оперативного состава органов управления, отвечающих за разработку и выдачу информационных продуктов, необходимых для лиц, принимающих решения по конкретной обстановке или ситуации в подконтрольной им сфере деятельности.

Информационно-аналитическая система состоит из приложений прикладного программного обеспечения, обеспечивающих решение задач ситуационного

управления: средств, реализующих анализ ситуаций и их моделирование, системы сбора, распознавания и хранения информации, информационного и лингвистического обеспечения (словарей и классификаторов), информационно-аналитического и информационно-расчетного программного обеспечения.

Программы в составе ИАС должны обеспечивать экспертно-аналитический и оперативный состав ИАСЦ пользовательскими средствами анализа обстановки, которые позволят выполнять [1]:

- оценку параметров ненаблюдаемых (скрытых) элементов обстановки на основе выборочных или косвенных данных по результатам мониторинга, выявление и извлечение фактов;
- оперативный анализ обстановки путем ее сравнения с прошедшим периодом (без изменений, хуже, лучше, аномалия) с целью выявления ситуаций, требующих немедленного реагирования;
- оценку ситуации с целью определения необходимости выработки решений по ее нормализации и степени сложности ситуации — штатная, критическая, чрезвычайная;
- прогнозирование изменения обстановки — без управляющего воздействия, с управляющим воздействием, сценарное прогнозирование с учетом внешних факторов;
- поддержку процессов принятия управленческих решений — адаптацию типовых решений и выработку нетиповых решений на основе экспертной поддержки (с учетом прогнозирования).

Информационно-аналитическая система — это ключевой и самый наукоемкий компонент ИАСЦ, процессы его проектирования и реализации должны стать приоритетными при реализации концептуального подхода к управлению процессами стратегического планирования в области национальной безопасности, обсуждаемого в [1] и в настоящей работе. Этот компонент заслуживает отдельного глубокого исследования, в результате которого должны быть детализированы все решаемые задачи, формально описаны используемые данные и выходные документы.

Вместе с тем аналитические инструменты — далеко не единственный вопрос технологического уровня ИАСЦ. Помимо прикладных программ ИАС создание ИАСЦ потребует самостоятельных архитектурных решений в отношении инфраструктуры ИАСЦ, решений в отношении технического, возможно даже аппаратного, обеспечения, общесистемного и специального программного обеспечения, средств телекоммуникаций и обеспечения информационной безопасности. Эти решения должны учитывать как описанную функциональность ИАСЦ, так и современные достижения и тенденции ИТ-индустрии, а также иные обстоятельства и факторы, которые окажут влияние на функционирование ИАСЦ. Изучение этих вопросов — предмет дальнейших исследований.

Литература

1. *Сучков А. П., Босов А. В., Макошко А. А.* Ситуационный анализ в процессах стратегического планирования в области обеспечения национальной безопасности // *Инновации*, 2018. Т. 238. Вып. 8. С. 33–39.
2. О формировании системы распределенных ситуационных центров, работающих по единому регламенту взаимодействия: Указ Президента РФ № 648 от 25.07.2013.
3. Концепция создания Системы распределенных ситуационных центров органов государственной власти Российской Федерации от 3 октября 2013 г., № Пр-2308.
4. Технологический портал СМЭВ. <http://smev.gosuslugi.ru>.
5. Единый портал государственных услуг. <https://www.gosuslugi.ru>.
6. *Нежданов И. Ю.* Технологии информационных войн в интернете // *Наша молодежь*, 2016. № 3. С. 28–31.

Поступила в редакцию 19.07.18

CONCEPTUAL APPROACHES TO CREATING AN INFORMATION AND ANALYTICAL SITUATIONAL CENTER FOR STRATEGIC PLANNING IN THE FIELD OF NATIONAL SECURITY

A. P. Suchkov and A. V. Bosov

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

Abstract: The article considers approaches to creation of an information and analytical situational center (IASC) for managing strategic planning processes in the field of national security. From the point of view of the study of the role of situational analysis in these processes, two complete and interconnected management cycles can be distinguished which are conditionally called “planning” and “realization.” The functional model of the IASC implementing these management cycles is substantiated, the composition and methods of analyzing information resources are outlined, focused on monitoring, detection, prediction, and assessment of threats. Data sources, potentially significant for the IASC monitoring subsystem, are considered.

Keywords: strategic planning; situational analysis; management cycle; information threat model; situational center; data sources

DOI: 10.14357/08696527180408

References

1. Suchkov, A. P., A. V. Bosov, and A. A. Makosko. 2018. Kontseptual’nye podkhody k sozdaniyu informatsionno-analiticheskogo situatsionnogo tsentra v interesakh strategicheskogo planirovaniya v oblasti obespecheniya natsional’noy bezopasnosti [Situational

analysis in the processes of strategic planning in the field of national security]. *Innovations* 238(8):33–39.

2. O formirovaniy sistemy raspredelennykh situatsionnykh tse ntrov, rabotayushchikh po edinomu reglamentu vzaimodeystviya [On the formation of a system of distributed situational centers operating under the single rules of interaction]. 2013. Ukaz Prezidenta RF No. 648 ot 25.07.2013 [Decree of the President of RF No. 648 dated 25.07.2013].
3. Kontseptsiya sozdaniya Sistemy raspredelennykh situatsionnykh tse ntrov organov gosudarstvennoy vlasti Rossiyskoy Federatsii [The concept of creating a system of distributed situational centers of the government authorities of the Russian Federation]. October 3, 2013. No. Pr-2308.
4. Tekhnologicheskiy portal SMEV [SMEV Technology Portal]. Available at: <http://smev.gosuslugi.ru> (accessed September 1, 2018).
5. Edinyy portal gosudarstvennykh uslug [Unified portal of public services]. Available at: <https://www.gosuslugi.ru> (accessed September 1, 2018).
6. Nezhdanov, I. Yu. 2016. Tekhnologii informatsionnykh voyn v Internete [Information wars on the Internet]. *Nasha molodezh'* [Our Youth] 3:28–31.

Received July 19, 2018

Contributors

Suchkov Alexander P. (b. 1954) — Doctor of Science 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 119333, Russian Federation; ASuchkov@ipiran.ru

Bosov Alexey V. (b. 1969) — Doctor of Science in technology, 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; AVBosov@ipiran.ru

ПРОЦЕССНАЯ МОДЕЛЬ МОДЕРНИЗАЦИИ И РАЗВИТИЯ ИНФОРМАЦИОННЫХ СИСТЕМ НА ВСЕХ СТАДИЯХ ЖИЗНЕННОГО ЦИКЛА

А. П. Сучков¹

Аннотация: Создание процессной модели модернизации и развития информационных систем (ИС) представляет собой сложную и актуальную научную проблему. Обсуждается определение стадий жизненного цикла (ЖЦ) ИС, для всех стадий выявлены процессы, ответственные за модернизацию и развитие системы, сформирована совокупность процессов ЖЦ, обеспечивающая возможность непрерывного развития и модернизации ИС в связи с возникновением новых функциональных требований, выявлением ошибок, с учетом новых предложений по техническим, организационным, кадровым и нормативно-правовым изменениям. Реализация указанной процессной модели в соответствии с ГОСТом должна осуществляться службой поддержки с использованием специализированной системы управления модернизацией и развитием (СУМР).

Ключевые слова: процессный подход; процессная модель; модернизация и развитие информационной системы; система управления

DOI: 10.14357/08696527180409

1 Введение

Требование обеспечения непрерывности модернизации и развития ИС в ходе ее создания и эксплуатации предполагает углубленный учет фактора времени с точки зрения обеспечения жизнеспособности системы, т. е. рассмотрения происходящих в ИС явлений как процессов, имеющих характеристики, изменяющиеся во времени. Реализовать системный подход в рамках описания таких явлений предлагается на основе применения *процессного подхода*.

Процесс определяется как совокупность взаимосвязанных и взаимодействующих видов деятельности, преобразующих входы в выходы (ИСО 9004:2000 «Системы менеджмента качества. Руководство по выполнению усовершенствований»). Рассматривая ИС как систему, развивающуюся во времени, необходимо использовать понятие *жизненного цикла* системы как сложноорганизованной совокупности процессов. Общие основы для описания ЖЦ систем, созданных

¹Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, ASuchkov@ipiran.ru

людьми, а также определение детально структурированных процессов и соответствующую терминологию устанавливает ГОСТ Р ИСО/МЭК 15288-2005 «Процессы жизненного цикла систем».

Для того чтобы охарактеризовать процессы обеспечения модернизации и развития ИС необходимо:

- определить все стадии ЖЦ ИС;
- для всех стадий выявить процессы ЖЦ ИС, ответственные за модернизацию и развитие системы;
- сформировать систему процессов ЖЦ, обеспечивающую возможность непрерывного развития и модернизации ИС в связи с возникновением новых функциональных требований, выявлением ошибок, с учетом новых предложений по техническим, организационным, кадровым и нормативно-правовым изменениям.

2 Основные стадии жизненного цикла информационной системы

В соответствии с ГОСТ Р ИСО/МЭК 15288-2005 можно выделить 6 основных стадий ЖЦ ИС (табл. 1 содержит краткую характеристику стадий):

Таблица 1 Пример стадий ЖЦ, их целей и основных схем решений

Стадия ЖЦ	Цель	Схема решений
Замысел	Определить потребности правообладателей Исследовать замыслы Предложить жизнеспособные решения	Вариант решения: – выполнить следующую стадию; – продолжить данную стадию; – вернуться к предыдущей стадии; – приостановить проект; – завершить проект
Разработка	Уточнить требования к системе Создать описание решений Создать систему Провести верификацию и валидацию системы	
Производство	Произвести систему Проконтролировать и испытать	
Применение	Обеспечить применение системы для удовлетворения потребностей пользователей	
Поддержка применения	Обеспечить модернизацию, развитие и устойчивую реализацию возможностей системы	
Прекращение применения и списание	Хранение, архивирование или списание системы	

- (1) стадия замысла;
- (2) стадия разработки;
- (3) стадия производства;
- (4) стадия применения;
- (5) стадия поддержки применения;
- (6) стадия прекращения применения и списания.

Таблица 2 содержит обобщенные данные о функционале модернизации и развития системы, выявленные в ходе анализа процессов ЖЦ ИС. В таблице упоминаются две организационные структуры, оговоренные в ГОСТ Р ИСО/МЭК 15288-2005:

- (1) *правообладатель* — сторона, имеющая право, долю или претензии на систему или на владение ее характеристиками, удовлетворяющими потребности и ожидания этой стороны;

Таблица 2 Функционал модернизации и развития ИС

Стадия ЖЦ	Функционал модернизации и развития	Объект модернизации и развития	Субъект модернизации и развития
Замысел	Формулировка стратегии развития Формулировка приемлемых требований по модернизации системы Формирование требований к службе поддержки	Базовая линия требований правообладателей и предварительных системных требований	Правообладатель
Разработка	Модернизация проектных решений «на лету»	Обратная связь между правообладателями и теми, кто будет производить, управлять, использовать, поддерживать и списывать рассматриваемую систему Оцененные и уточненные системные требования, бюджет проекта, базовые сроки выполнения и оценки затрат для собственника ЖЦ	Служба поддержки
Производство	Мониторинг хода реализации базовой линии требований правообладателей и системных требований, выявление недостатков и несоответствий в ходе производства	Проектные решения (путем возврата к стадии разработки)	Служба поддержки

Окончание табл. 2 на с. 89

Таблица 2 (окончание) Функционал модернизации и развития ИС

Стадия ЖЦ	Функционал модернизации и развития	Объект модернизации и развития	Субъект модернизации и развития
Применение	<p>Мониторинг хода реализации стратегии развития</p> <p>Мониторинг характеристик функционирования, идентификация, классификация и составление отчетов об отклонениях, недостатках и отказах</p> <p>Мониторинг ресурсообеспеченности (персонал, материально-техническое обеспечение)</p> <p>Идентификация проблем или недостатков, информирование соответствующих организаций (пользователи, разработчики, производители или обслуживающие органы) о необходимости проведения корректирующих действий</p> <p>Выявляются и анализируются новые возможности для совершенствования функционала рассматриваемой системы через обратную связь с правообладателем</p> <p>Иницируются корректирующие действия по устранению ранее не обнаруженных конструкционных ошибок</p> <p>Появление различных модификаций системы</p>	<p>Информационная система:</p> <ul style="list-style-type: none"> – техническое обслуживание; – незначительная (с низкой стоимостью и кратковременная) модификация (относится к стадии поддержки применения); – значительная (постоянная) модификация и продление срока жизни рассматриваемой системы (относится к стадии разработки производства); – изъятие и списание при окончании срока жизни (относится к стадии изъятия и списания) 	Служба поддержки
Поддержка применения	<p>Контроль рабочих характеристик служб и системы поддержки, идентификация, классификация и составление отчетов об аномалиях, недостатках и отказах служб и системы поддержки</p>	<p>Служба поддержки:</p> <ul style="list-style-type: none"> – техническое обслуживание; – незначительные модификации служб и системы поддержки; – существенные модификации служб или системы поддержки (относится к стадии разработки и производства) 	Правообладатель
	<p>Исправляются недостатки, допущенные при проектировании системы</p>	Проектные решения	Служба поддержки
Прекращение применения и списание	Нет функционала модернизации		

- (2) *служба поддержки* — организационная структура, создаваемая правообладателями и отвечающая за поддержку и развитие системы на всех стадиях ЖЦ.

3 Процессная модель реализации стратегии модернизации и развития информационной системы на всех стадиях ее жизненного цикла

Эффективным способом обеспечения непрерывного поиска, разработки и оперативного внедрения в ИС новых функциональных возможностей и исправления ошибок является создание *системы управления модернизацией и развитием информационной системы*, позволяющей в полной мере реализовать определенный выше функционал.

В [1, 2] обоснована циклическая функциональная модель систем ситуационного управления, включающая 5 стадий: целеполагание, мониторинг, анализ, решение и действие. В развитие этого подхода и предлагается модель СУМР ИС, ориентированная на процессный подход в соответствии с упомянутым стандартом.

Рассмотрим содержание этого функционала с точки зрения пятистадийной модели системы управления (рис. 1).

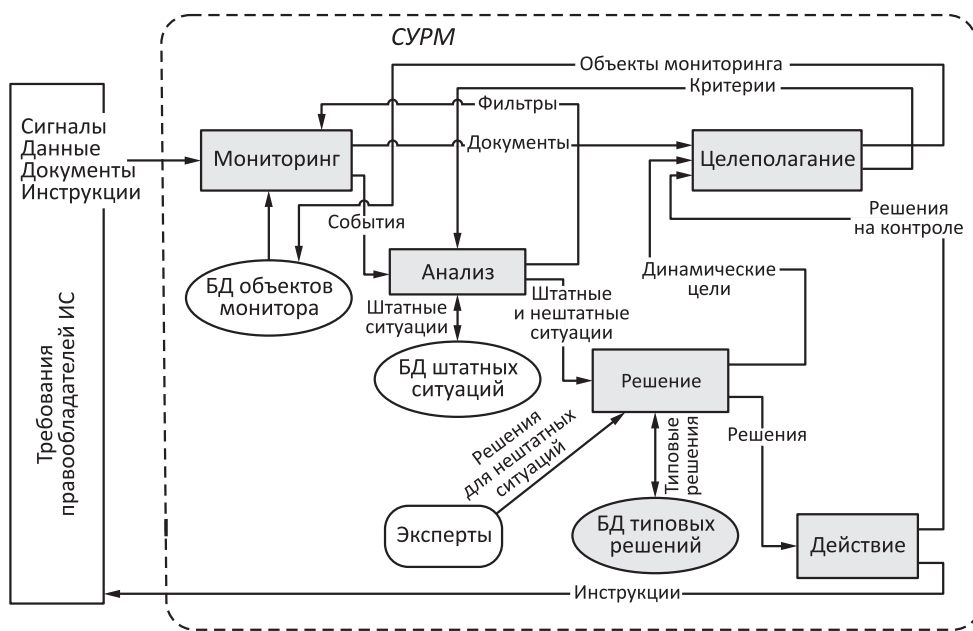


Рис. 1 Функциональная модель СУМР

Целеполагание. Функционирование системы управления является целесообразным и определяется нормативно-правовыми документами, регламентирующими процессы ее функционирования. Система целей обычно формализуется в виде целевых показателей путем установления количественных и временных критериев их целевых значений (планирование) [3]. Кроме того, при ситуационном управлении возможно установление динамических целей и корректировка стратегических и тактических целей. Целеполагание обычно осуществляется организацией — разработчиком замысла системы на основе общих требований и соглашений с правообладателем. Состав и критерии целевых показателей устанавливаются на стадии замысла в процессе определения и анализа требований правообладателей и формирования стратегии модернизации и развития и могут быть откорректированы на последующих стадиях ЖЦ с использованием процессов соглашения, верификации, валидации и контроля качества.

Мониторинг. На данной стадии в рамках СУМР обеспечивается сбор событийной информации:

- о состоянии значений целевых показателей и динамических целей системы;
- о характеристиках функционирования ИС и вспомогательных систем;
- о возможных изменениях требований правообладателей по отдельным функциональным возможностям ИС.

На основе этого осуществляется идентификация, классификация и составление отчетов о ходе реализации стратегии модернизации и развития ИС, об отклонениях, недостатках и отказах.

Мониторинг опирается на сенсорную составляющую системы управления, включая объективные и субъективные средства наблюдения, а также различные сканеры информационного пространства. Он осуществляется средствами СУМР на всех стадиях ЖЦ с использованием процессов функционирования, управления информацией, управления качеством, оценки и контроля проектов, управления конфигурацией, управления ресурсами, управления рисками.

Анализ. Осуществляется оценка сложившейся ситуации (штатная, нештатная) и формирование гипотез о состоянии ИС, а также анализ и учет существующего опыта. Стадия опирается на методы декомпозиции складывающихся ситуаций на известные или достаточно простые ситуации для нормализации с помощью ресурсов системы. В зависимости от того, какая ситуация идентифицирована, на обнаруженные проблемы вырабатываются варианты реакции СУМР:

- отсутствие действия;
- техническое обслуживание (стадия применения);
- незначительная (с низкой стоимостью и кратковременная) модификация (относится к стадии поддержки применения);

- значительная (постоянная) модификация и продление срока жизни рассматриваемой системы (относится к стадии разработки и производства, а иногда и к стадии замысла);
- изъятие и списание при окончании срока жизни (относится к стадии изъятия и списания).

Анализ данных мониторинга осуществляется на всех стадиях ЖЦ с использованием процессов планирования проекта, оценки проекта, принятия решений, анализа требований, определения требований, функционирования, обслуживания, верификации и валидации.

Решение. На основании процессов управления ресурсами и принятия решений осуществляется сравнительный анализ вариантов решения и выбор одного из них. Выбор решения осуществляется на основании следующих критериев:

- влияние на целевые показатели;
- ресурсные ограничения;
- оценки «стоимости эффективности».

С точки зрения СУМР решение осуществляется в отношении:

- обслуживания системы и продолжения стадии применения;
- перехода на стадию разработки и производства с учетом новых требований правообладателей по расширению функциональности системы;
- перехода на стадию замысла при необходимости существенной модернизации и корректировки целевых показателей.

Принятое решения обуславливает формирование динамических целей СУМР для возможности контроля за его исполнением.

Действие. На данной стадии управления осуществляется выполнение выработанного решения с учетом применяемых управляющих воздействий, исходя из имеющихся ресурсов, путем активизации соответствующих процессов ЖЦ ИС.

Пять стадий управления реализуют основной цикл управления в рамках СУМР с обратной связью, так как после применения управляющих воздействий осуществляется переход к стадии мониторинга для контроля исполнения и, возможно, к стадии целеполагания для установления и корректировки динамических целей системы (рис. 2).

Итак, с точки зрения функционала СУМР определено содержание всех стадий ЖЦ ИС по ГОСТ Р ИСО/МЭК 15288-2005, выделены все его положения о задачах модернизации и развития, а также то, как этот функционал укладывается в стандартную модель системы управления. Теперь можно перейти к формированию процессной модели реализации стратегии модернизации и развития ИС на всех стадиях ее ЖЦ на основе систематического применения процессов ЖЦ.

Для обеспечения унифицированного описания процессов согласно ГОСТ Р 57098-2016/ISO/IEC TR 24774 2010 «Управление жизненным циклом. Руководство для описания процесса» используются следующие элементы процесса:

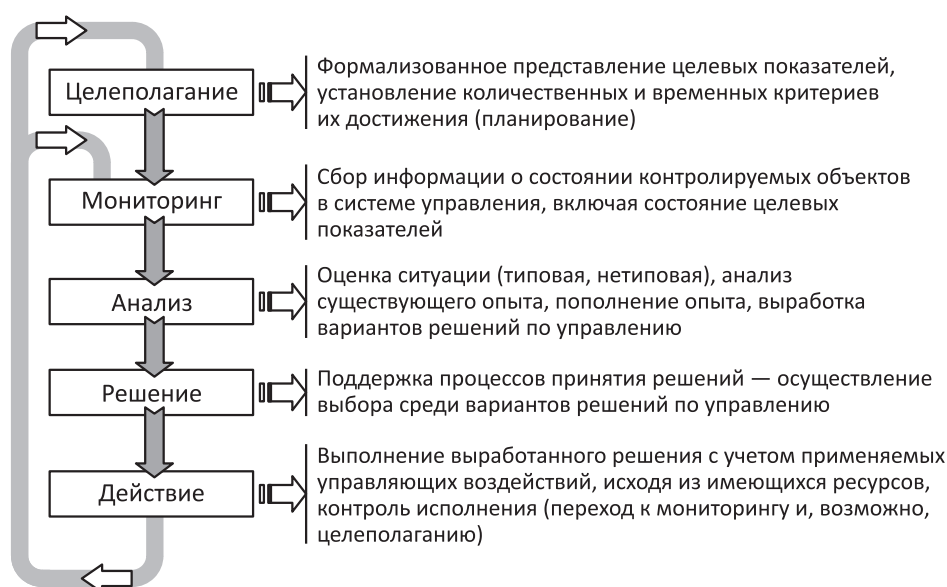


Рис. 2 Структура основного цикла управления

- **название** — описывает назначение процесса;
- **цель** — описывает цель выполнения процесса;
- **выходы** (выходные результаты) — выражают заметные результаты, ожидаемые от успешного выполнения процесса;
- **действия** — перечень действий, которые могут использоваться для достижения выходных результатов. Каждое действие может быть в дальнейшем переработано в группу действий более низкого уровня;
- **задачи** — специальные действия, которые могут осуществляться для выполнения какого-то действия. Множество взаимосвязанных задач часто объединяется в пределах какого-то действия;
- **информационные объекты** — отдельно идентифицируемые содержательные части информации, произведенные и сохраненные для использования человеком в течение жизненного цикла системы или программных средств.

На рис. 3 изображена обобщенная схема процессной модели системы модернизации и развития, представленная в двух измерениях: стадии ЖЦ ИС и стадии цикла управления с учетом функционала СУМР и применения процессной модели системы управления [2].

В соответствии со схемой процессной модели СУМР должна реализовывать следующий функционал:

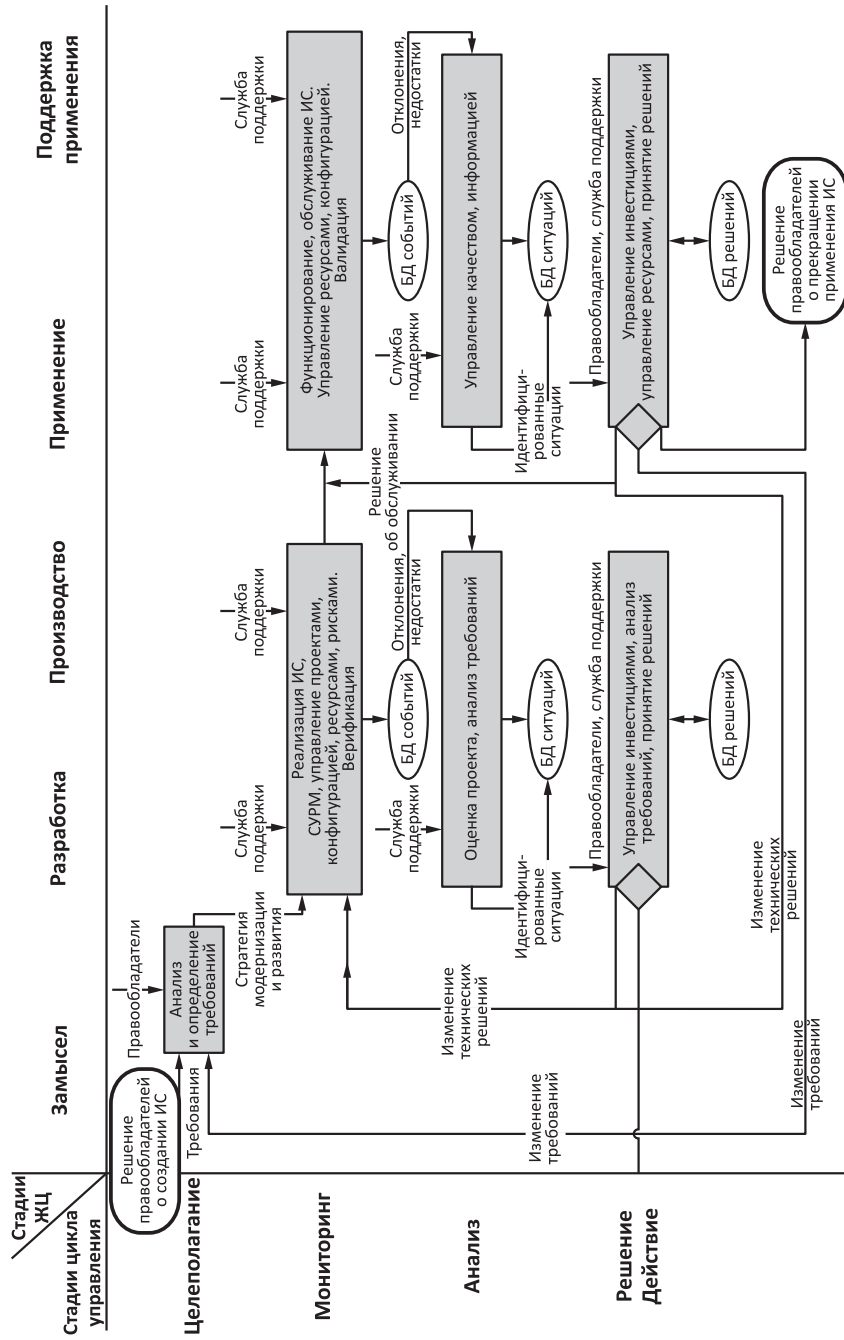


Рис. 3 Процессная модель СУМР на всех этапах ЖЦ ИС

- мониторинг хода реализации базовой линии требований правообладателей и системных требований, выявление недостатков и несоответствий в ходе разработки и производства ИС;
- мониторинг хода реализации стратегии развития;
- мониторинг характеристик функционирования, идентификация, классификация и составление отчетов об отклонениях, недостатках и отказах;
- мониторинг ресурсообеспеченности (персонал, материально-техническое обеспечение);
- идентификация проблем или недостатков, информирование соответствующих организаций (пользователи, разработчики, производители или обслуживающие органы) о необходимости проведения корректирующих действий;
- выявление и анализ новых возможностей для совершенствования функционала ИС через обратную связь с правообладателем;
- инициирование корректирующих действий по устранению ранее не обнаруженных конструктивных ошибок;
- техническое обслуживание ИС (на стадии применения);
- незначительная (с низкой стоимостью и кратковременная) модификация (относится к стадии поддержки применения);
- значительная (постоянная) модификация и продление срока жизни рассматриваемой системы (относится к стадии разработки и производства, а иногда и к стадии замысла);
- контроль рабочих характеристик служб и системы поддержки, идентификация, классификация и составление отчетов об аномалиях, недостатках и отказах служб и системы поддержки;
- изъятие и списание ИС по окончании срока жизни (относится к стадии изъятия и списания).

Для каждого процесса (группы процессов) определены входы и выходы, субъект осуществления функционала процесса. Как можно видеть, процессная модель СУМР содержит определенные информационные ресурсы, формируемые в процессе создания и эксплуатации. Сюда относятся:

- база данных (БД) событий, которая должна содержать сведения обо всех изменениях состояния объектов мониторинга (имеющих как позитивную, так и негативную окраску с точки зрения функционирования ИС);
- БД ситуаций, которая должна содержать идентифицированные ситуации, сложившиеся в отношении эксплуатируемого банка данных и являющиеся совокупностью регистрируемых взаимосвязанных причинно-следственными связями событий;

- БД решений, которая предназначена для накопления знаний в процессе эксплуатации банка данных и обеспечения их учета, хранения и распространения с целью аудита и сохранения этих знаний в системе в соответствии с ГОСТ Р ИСО/МЭК 15288-2005.

Приведенная модель имеет концептуальный, обобщенный характер, в ней объединены в функциональные блоки несколько процессов, обладающих в соответствии с ГОСТ Р ИСО/МЭК 15288-2005 определенным функционалом модернизации и развития.

Литература

1. Зацаринный А. А., Сучков А. П. Системотехнические подходы к созданию системы поддержки принятия решений на основе ситуационного анализа // Информатика и её применения, 2016. Т. 10. Вып. 4. С. 105–113.
2. Сучков А. П. Полнофункциональный процессный подход к реализации систем ситуационного управления // Системы и средства информатики, 2017. Т. 27. № 1. С. 86–100.
3. Сучков А. П. Формирование системы целей для ситуационного управления // Системы и средства информатики, 2013. Т. 23. № 2. С. 171–182.

Поступила в редакцию 30.05.18

PROCESS MODEL OF INFORMATION SYSTEMS MODERNIZATION AND DEVELOPMENT AT ALL STAGES OF THEIR LIFE CYCLE

A. P. Suchkov

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

Abstract: Creating a process model of modernization and development of information systems (IS) is a complex and urgent scientific problem. The article discusses the definition of stages of the life cycle (LC) of IS. For all stages the processes responsible for modernization and development of the system are identified. A set of processes of LC is formed, providing the possibility of continuous development and modernization of IS in connection with the emergence of new functional requirements, errors, taking into account new proposals for technical, organizational, personnel, and regulatory changes. Implementation of this process model in accordance with the ISO standard should be carried out by a support service using a specialized management system of modernization and development.

Keywords: process approach; process model; information system modernization and development; management system

DOI: 10.14357/08696527180409

References

1. Zatsarinny, A. A., and A. P. Suchkov. 2016. Sistemotekhnicheskie podkhody k sozdaniyu sistemy podderzhki prinyatiya resheniy na osnove situatsionnogo analiza [Systems engineering approaches to a decision support system based on situational analysis]. *Informatika i ee Primeneniya — Inform. Appl.* 10(4):105–113.
2. Suchkov, A. P. 2017. Polnofunktional'nyy protsessnyy podkhod k realizatsii sistem situatsionnogo upravleniya [A fully functional process-based approach to the implementation of systems of situational management]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 27(1):86–100.
3. Suchkov, A. P. 2013. Formirovanie sistemy tseley dlya situatsionnogo upravleniya [The formation of the objective system to situational management]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 23(2):171–182.

Received May 30, 2018

Contributor

Suchkov Alexander P. (b. 1954) — Doctor of Science 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 119333, Russian Federation; ASuchkov@frccsc.ru

НЕКОТОРЫЕ СИСТЕМОТЕХНИЧЕСКИЕ ВОПРОСЫ ПРЕДОСТАВЛЕНИЯ ВЫЧИСЛИТЕЛЬНЫХ РЕСУРСОВ ДЛЯ НАУЧНЫХ ИССЛЕДОВАНИЙ В ГИБРИДНОЙ ВЫСОКОПРОИЗВОДИТЕЛЬНОЙ ОБЛАЧНОЙ СРЕДЕ*

*К. И. Волович*¹

Аннотация: Рассматриваются подходы к предоставлению научным коллективам вычислительных ресурсов для решения прикладных и фундаментальных научных задач. Предлагаются методы и алгоритмы адаптации и исполнения прикладных задач пользователей на гибридном вычислительном кластере, используемом для выполнения вычислений специализированные ускорители. Предлагаются решения по адаптации программного кода на языках высокого уровня к целевой платформе гибридного вычислительного кластера, алгоритмы предоставления задачам вычислительных ресурсов в условиях многозадачной среды, а также методы создания адаптированной среды исполнения заданий с использованием технологий виртуализации.

Ключевые слова: высокопроизводительные вычисления; кластер; виртуализация; управление вычислительным процессом; распределенные вычисления; параллельные вычисления; графический ускоритель; НРС

DOI: 10.14357/08696527180410

1 Введение

В настоящее время наблюдается тенденция предоставления вычислительных ресурсов для решения фундаментальных и прикладных задач в виде облачных сервисов [1–3]. В рамках классических подходов к облачным вычислениям такие сервисы могут быть предоставлены как программное обеспечение — SaaS (software-as-a-service), платформа — PaaS (platform-as-a-service) или инфраструктура — IaaS (infrastructure-as-a-service). В [4, 5] предложен также вариант предоставления научного исследования как сервис — RaaS (research-as-a-service).

Такой же облачный подход целесообразно применять и к услугам высокопроизводительных вычислений, выполняемых в интересах научных коллективов различных направлений прикладных и фундаментальных наук. Это позволит централизовать ресурсы и обеспечить загрузку вычислительных комплексов расчетными задачами с большей эффективностью, чем при монопольном предоставлении каждому научному коллективу вычислительных мощностей, а также позволит обеспечить большую гибкость в эксплуатации вычислительного комплекса, сократить время выполнения вычислений [6, 7].

* Работа выполнена при частичной поддержке РФФИ (проекты 18-07-00669, 18-29-03100).

¹ Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, KVolovich@frccsc.ru

Однако при предоставлении услуг высокопроизводительных вычислений необходимо учитывать специфику организации вычислительного процесса и архитектуру высокопроизводительного комплекса.

Современные вычислительные комплексы, предназначенные для решения научных задач, строятся, как правило, на основе гибридных архитектур, включающих в себя как центральные процессоры общего применения, так и специализированные ускорители математических операций. Так, по состоянию на июнь 2018 г. восемь из десяти первых суперкомпьютеров из рейтинга Top 500 имеют гибридную архитектуру [8], в которую могут входить центральные процессоры (CPU) различных архитектур (например, Intel x86_64, IBM Power), а также ускорители вычислений (GPU, graphics processing unit) различных архитектур (например, Intel Xeon Phi, Nvidia Tesla, Matrix coprocessor).

Для использования возможностей вычислений на таких комплексах в рамках концепции облачных вычислений необходимо применять определенные методики организации вычислительного процесса, обеспечивающие разделение ресурсов графических ускорителей между клиентами вычислительного комплекса.

Кроме того, необходимо учитывать, что любой научный коллектив, использующий для решения своих задач гибридный вычислительный комплекс, как правило, имеет программный задел, предназначенный для решения его научных задач. В общем случае такой программный задел может быть не совместим с архитектурой и программной средой гибридного вычислительного комплекса. Вопросы классификации прикладного программного обеспечения по совместимости и мобильности рассмотрены в [9].

В данной статье предлагаются системотехнические подходы к проблемам адаптации программного кода научных коллективов к среде гибридного высокопроизводительного комплекса, создания индивидуальной среды исполнения задач с использованием технологии виртуализации на основе программных контейнеров, а также разработки алгоритма управления ресурсами вычислительного комплекса [10, 11].

2 Адаптация программного кода к архитектуре гибридного высокопроизводительного комплекса

Имеющийся у научных коллективов задел программного обеспечения, предназначенного для решения фундаментальных и прикладных научных задач, не может выполняться на средствах гибридного высокопроизводительного комплекса без проведения определенной подготовки.

Для адаптации имеющегося у научного коллектива задела по решению прикладной или фундаментальной задачи в составе гибридного высокопроизводительного комплекса необходимо создать ряд инструментальных средств, позволяющих выполнить загрузку, обработку, тестирование и отладку и, наконец, исполнение программного кода с целью решения научной задачи. Финальной

стадией этого процесса является оценка улучшения либо деградации производительности при решении данной задачи по сравнению с ранее использованными вычислительными комплексами.

Начальным этапом адаптации программного кода к условиям выполнения на средствах гибридного высокопроизводительного комплекса является классификация программного кода по степени мобильности [9].

Такая классификация может быть выполнена как непосредственно пользователем, так и с использованием средств автоматизации. В частности, в качестве такого инструментального средства может быть использована нейронная сеть, обученная на предварительно размеченном наборе данных.

Пользовательская классификация, на первый взгляд, является более чем достаточной для принятия решения, к какому классу относится адаптируемая задача. В то же время необходимо учитывать скрытые нюансы, которые могут быть недоступны пользователю в момент проведения работ.

С целью учета такой скрытой информации предлагается использовать компоненты искусственного интеллекта — нейронную сеть, предварительно обучающуюся на стартовом наборе размеченных исходных данных и в дальнейшем продолжающую свое обучение в процессе работы системы. Таким образом, на начальном (пусковом) этапе нейронная сеть проходит обучение на стартовом массиве данных и корректирует свое состояние на базе информации, получаемой в процессе работы пользователей по адаптации своих научных задач для выполнения на гибридном высокопроизводительном комплексе.

На рис. 1 представлен сценарий ввода в эксплуатацию и функционирования нейронной сети классификации программного кода и корректировки ее состояния на основе данных, получаемых при работе пользователей по адаптации программного кода.

Отметим, что набор данных для первичного обучения нейронной сети может быть построен как на основе реальных примеров программного кода, так и на основе искусственно сгенерированных (синтетических) примеров. Разметка таких примеров для первичного обучения нейронной сети в любом случае должна выполняться человеком.

Набор данных для первичного обучения должен включать примеры для всех языков программирования высокого уровня (ЯВУ), применяемых для разработки прикладного кода в гибридном высокопроизводительном комплексе (например, Python, TensorFlow). Однако архитектура блока классификации может быть различной.

Возможен вариант обучения единой нейронной сети на полном наборе данных первичного обучения, включающем примеры на различных языках программирования (сеть «НС Е» на рис. 1).

Другой вариант построения блока классификации предполагает наличие группы нейронных сетей, каждая из которых выполняет классификацию программного кода для одного языка программирования. На рис. 1 показаны нейронные сети «НС 1»–«НС 3» для каждого языка высокого уровня.

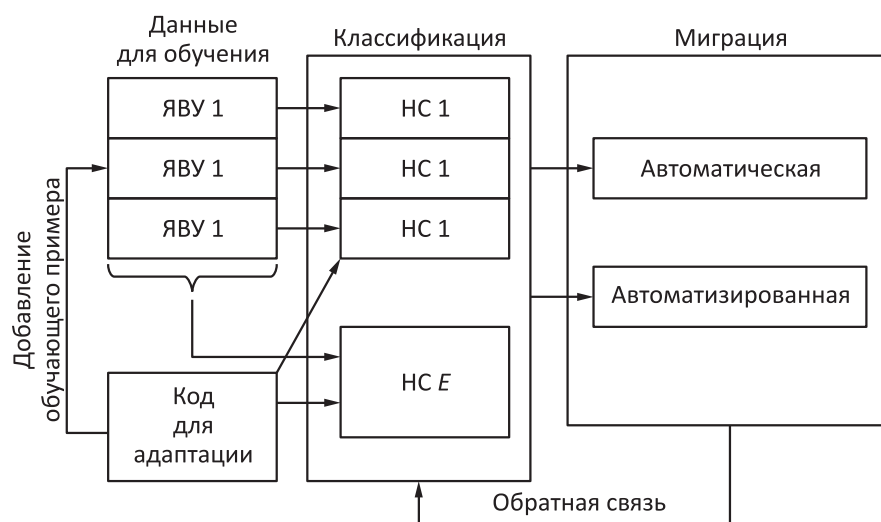


Рис. 1 Сценарий ввода в эксплуатацию и функционирования нейронной сети

С точки зрения универсальности и построения полностью автоматического модуля классификации программного кода первый вариант предпочтительнее, поскольку не требует предварительной сортировки программного кода на основе используемого языка. Кроме того, такой подход позволяет проводить классификацию текстов программ, разработанных с использованием нескольких языков программирования, что является обычной практикой при создании больших проектов. Недостатки такого варианта — длительность обучения и высокие требования к затрачиваемым ресурсам.

Вариант с разделением блока классификации на компоненты анализа одного выбранного языка программирования представляется более целесообразным для построения макетов и отработки технологии в силу меньшей требовательности к вычислительным ресурсам.

Помимо стартового набора исходных программных текстов для обучения нейросети в процессе работы используются тексты, прошедшие классификацию и дальнейшие шаги по адаптации программного кода к условиям выполнения на средствах высокопроизводительного комплекса. Входными данными для такого вторичного обучения служит сам программный текст (на рис. 1 — код для адаптации), а также данные о дальнейших шагах по адаптации (обратная связь), при этом программный текст добавляется к исходному набору обучающих примеров в качестве дополнительного размеченного компонента.

Бизнес-процесс по выполнению адаптации имеющегося программного кода и последующего его выполнения на средствах гибридного высокопроизводительного кластера включает следующие этапы [9]:

- (1) классификация приложений;
- (2) подготовка исполняемого кода;
- (3) создание заданий для вычислителей;
- (4) организация и обслуживание очереди заданий;
- (5) подготовка вычислительной среды;
- (6) выполнение расчета;
- (7) предоставление результатов расчета.

Адаптация программного кода выполняется на этапах 1 и 2. В ходе выполнения этих этапов должна быть выполнена как подготовка самого исполняемого кода, так и подготовка описания необходимой вычислительной среды для исполнения этого кода. Сама среда по данному описанию формируется на этапе 5 и будет рассмотрена ниже.

Для адаптации компилируемого программного кода следует выполнить следующую последовательность действий в зависимости от результатов классификации, выполненных на первом этапе.

Для приложений, независимых от архитектуры, необходимо определить:

- перечень модулей и программных компонентов, сред исполнения для поддержки языка приложения (например, Python, TensorFlow);
- перечень системных библиотек, необходимых для поддержки программных компонентов прикладного программного обеспечения и среды;
- перечень необходимых драйверов устройств и программных компонентов (например, драйверы GPU и CUDA).

Поскольку приложения не зависят от вычислительной платформы, действия по их адаптации сводятся к определению перечней указанных компонентов и формированию на их основе необходимой среды исполнения с использованием механизмов виртуализации (см. ниже).

Для приложений, мобильных в рамках вычислительных архитектур, необходимо:

- осуществить компиляцию исходного кода в код целевой архитектуры;
- произвести компоновку целевого кода с системными и прикладными программными библиотеками;
- определить состав необходимых программных модулей и системных библиотек для формирования среды исполнения адаптированного кода.

Отметим, что для обоих типов приложений одним из результатов процесса адаптации является описание состава среды исполнения в виде набора библиотек, программных модулей и драйверов. Данное описание используется для создания среды исполнения пользовательского приложения в облачной инфраструктуре.

Для приложений, зависящих от архитектуры, адаптация к средствам гибридного вычислительного комплекса определяется индивидуальным подходом и сводится к замене зависимого от архитектуры кода его функциональным аналогом в рамках целевой архитектуры. Фактически такая работа является новой разработкой программного кода или его компонентов и не может быть отнесена к адаптации. Поэтому вопросы адаптации таких приложений к условиям работы на гибридной высокопроизводительной вычислительной системе выходят за рамки настоящей статьи.

Действия по компиляции, компоновке и составлению реестров необходимых программных библиотек и компонентов могут выполняться как в автоматическом, так и в автоматизированном режиме с участием человека.

Автоматический режим предпочтительно использовать для адаптации задач, использующих общепотребительное программное обеспечение, для которого имеется описание зависимостей программных модулей и библиотек. В этом случае инструментальные средства среды исполнения (например, интерпретаторы языков высокого уровня) имеют возможности построения дерева требуемых компонентов и автоматической генерации полного списка зависимостей.

В случае автоматизированной (с участием человека) компиляции и компоновки приложений, когда отсутствует точное описание состава программных модулей и зависимостей, действия по подбору необходимых программных библиотек осуществляются разработчиком в процессе интерактивного процесса сборки, компоновки и отладки.

В обоих случаях создается исполняемый код на целевом языке исполняющей системы, а также описание полного состава среды окружения (программные модули и библиотеки). В ходе исполнения пользовательского приложения на основе этого описания средствами виртуализации создается индивидуальная среда исполнения приложения.

3 Создание виртуальной среды функционирования заданий

Предпочтительным режимом функционирования гибридного вычислительного кластера в многопользовательском режиме является параллельное исполнение пользовательских научных задач. Целесообразность применения такого режима объясняется тем, что комплекс обладает разнородными вычислительными ресурсами, доступными для пользовательских приложений, однако сами приложения, как правило, с максимальной эффективностью используют только один тип ресурса, другие при этом нагружены слабо или вообще простаивают. Такая картина наблюдается при использовании для решения научных задач центральных процессоров и ускорителей вычислений. Как правило, алгоритмы прикладных задач рассчитаны на параллельные вычисления только на CPU либо GPU и одновременная полная нагрузка на оба типа процессоров не достигается. Поэтому имеет смысл организовывать вычислительный процесс так, чтобы обеспечить параллельное исполнение пользователь-

ских задач, которое приводит к утилизации всех ресурсов гибридного комплекса.

Как было показано выше, для каждого пользовательского приложения создается полный перечень программных модулей и библиотек среды исполнения. Необходимо отметить, что состав модулей, необходимых для функционирования различных пользовательских приложений, может быть несовместимым между собой, что делает невозможным функционирование нескольких пользовательских приложений в рамках одной среды исполнения.

Одновременное формирование и исполнение группы таких сред в рамках одного вычислительного комплекса может осуществляться с использованием средств виртуализации. В этом случае для каждого приложения формируется своя вычислительная среда, в рамках которой обеспечивается изоляция программных модулей и библиотек.

Если рассматривать различные подходы к виртуализации, то можно отметить, что наименее ресурсоемкой технологией является технология контейнеров, позволяющая создать среду окружения для прикладного процесса по файлу-описанию.

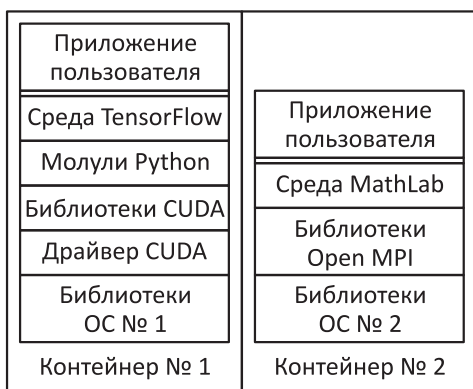


Рис. 2 Примеры контейнеров для выполнения пользовательских приложений

Фактически контейнер может быть создан динамически при загрузке задания на исполнение. Основой для файла-описания контейнера служит описание набора программных модулей, полученное в ходе работы по адаптации приложения.

На основе описания формируется структура исполняющей среды, включающая в себя последовательно: базовую операционную систему (ОС), специализированные драйверы устройств, интерфейсные библиотеки программных компонен-

тов параллельных вычислений, интегрированные среды разработки и исполнения программ, специализированные программные пакеты для прикладных научных исследований. На рис. 2 показан пример контейнеров, включающих в себя различные стеки программного обеспечения и предназначенные для исполнения разных прикладных задач.

На ресурсах гибридного вычислительного комплекса могут запускаться на исполнение по несколько экземпляров контейнеров каждого типа, обеспечивая тем самым необходимую загрузку ресурсов. При этом необходимо отметить, что для параллельного исполнения нескольких экземпляров одной задачи в рамках выполнения научного расчета код задачи должен быть адаптирован к условиям параллельной обработки. Например, использовать интерфейс MPI

(message passing interface) или другие механизмы реализации параллельных вычислений.

4 Алгоритмы управления ресурсами гибридного высокопроизводительного комплекса

Для обеспечения параллельного выполнения различных приложений с использованием виртуализации на базе технологии контейнеров в составе комплекса должна функционировать система управления вычислительными заданиями, обеспечивающая реализацию очередей заданий и политик обслуживания с учетом используемых вычислительных ресурсов.

Наличие в комплексе таких вычислительных ресурсов, как графические ускорители, накладывает на функционирование системы управления заданиями дополнительные требования по сравнению с классической архитектурой на CPU. Графический ускоритель выделяется задаче на определенный слот времени монопольно, а переключение между задачами, использующими GPU, является процессом долгим и снижает производительность комплекса.

Другой особенностью работы гибридного комплекса является то, что на его ресурсах должны выполняться одновременно задачи, производящие реальные

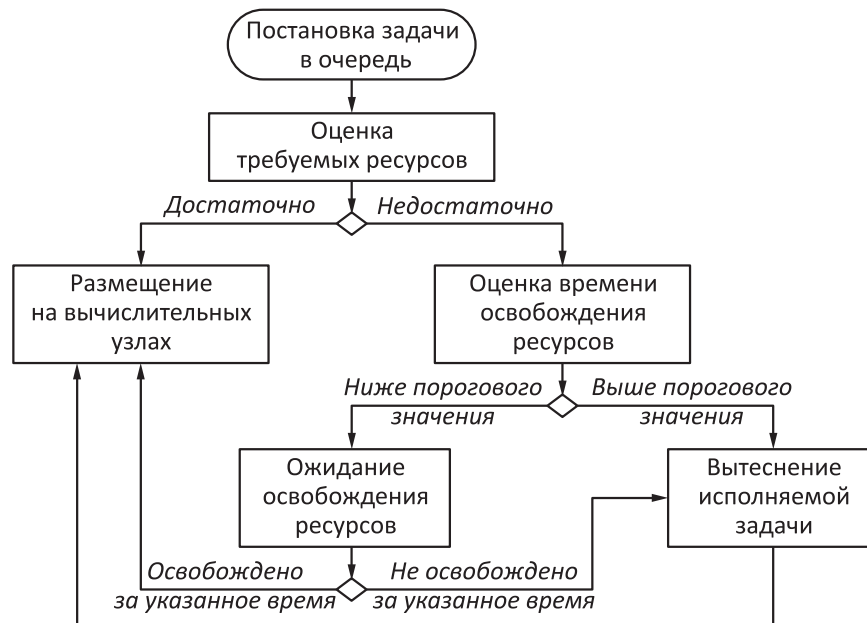


Рис. 3 Алгоритм управления ресурсами гибридного комплекса

расчеты и короткие тестовые задачи, необходимые для отладки и тестирования кода, но также требующие ресурсов GPU. Такие задачи требуют более высокого приоритета на исполнение, поскольку функционируют практически в интерактивном режиме с пользователями комплекса.

Исходя из перечисленных условий, предлагается алгоритм управления ресурсами гибридного комплекса, представленный на рис. 3.

Ключевым моментом данной процедуры является оценка предполагаемого времени исполнения. Первичная оценка может базироваться на заявленном пользователем предполагаемом времени исполнения программы, однако в дополнение к такой оценке полезно использовать статистические данные о времени запусков контейнера конкретной задачи в режиме реальных расчетов и тестовых, отладочных запусков.

В случае запуска нескольких экземпляров контейнера задачи каждому из них монополюно выделяются разные GPU во избежание переключений между разными экземплярами контейнера одного типа.

Следование такому алгоритму определения доступности ресурсов в сочетании с политиками обслуживания очередей с различными приоритетами позволит обеспечивать наиболее полную загрузку как центральных процессоров, так и ускорителей вычислений. Также обеспечивается приемлемое время реакции системы на запуск коротких задач тестирования и отладки. Одновременно с этим производится выполнение задач, выполняющих реальные расчеты.

5 Заключение

В статье рассмотрены вопросы подготовки адаптации программного обеспечения, предназначенного для научных расчетов, к функционированию на средствах гибридного высокопроизводительного вычислительного комплекса, а также создания виртуальной среды исполнения задач и управления ресурсами данного комплекса.

Предложено для автоматизированной адаптации программного кода использовать сценарии, включающие классификации программного кода по степени мобильности с использованием предварительно обученной нейронной сети. На основе проведенной классификации предлагаются сценарии автоматической и автоматизированной адаптации программного кода к условиям функционирования на высокопроизводительном вычислительном комплексе.

Разработаны подходы к созданию индивидуально настроенной среды исполнения прикладных задач с использованием механизма виртуализации вычислений на базе технологии контейнеров. Контейнеры создаются динамически в момент загрузки задач на вычисление по описаниям, полученным на этапе адаптации. Предложены методы и алгоритм управления ресурсами вычислительного комплекса, включая графические ускорители, в условиях одновременного параллельного исполнения прикладных научных задач различного типа, а также задач тестирования и отладки.

Рассмотренные в статье подходы и алгоритм могут быть использованы как основа для построения высокопроизводительного гибридного вычислительного комплекса, предоставляющего научным коллективам в качестве облачного сервиса инструментальные средства для адаптации программного кода к условиям данного комплекса, индивидуальную виртуальную среду исполнения и вычислительные ресурсы для проведения научных расчетов.

Литература

1. Wu W., Zhang H., Li Zh., Mao Ya. Creating a cloud-based life science gateway // e-Science and the Archaeological Frontier: 7th Conference (International) on eScience. — Piscataway, NJ, USA: IEEE, 2011. P. 55–61.
2. Ding F., Mey D., Wienke S., Zhang R, Li L. A study on today's cloud environments for HPC applications // 3rd Conference (International) on Cloud Computing and Services Science Proceedings. — Berlin, Germany: Springer, 2014. P. 114–127.
3. Volkov S., Sukhoroslov O. Simplifying the use of clouds for scientific computing with Everest // Procedia Comput. Sci., 2017. Vol. 119. P. 112–120.
4. Волович К. И., Зацаринный А. А., Кондрашев В. А., Шабанов А. П. О некоторых подходах к представлению научных исследований как облачного сервиса // Системы и средства информатики, 2017. Т. 27. № 1. С. 73–84.
5. Зацаринный А. А., Горшенин А. К., Волович К. И., Кондрашев В. А. Система управления научными сервисами как базовый сервис цифровой платформы для научных исследований // Проблемы управления научными исследованиями и разработками: Тр. 3-й научн.-практич. конф. — М.: ИПУ РАН, НИЦ «Ин-т им. Н. Е. Жуковского», 2017. С. 53–64.
6. Исследование вопросов управления результатами научно-исследовательской деятельности организаций, подведомственных ФАНО России, и научными сервисами сети ЦКП ФАНО: Отчет о НИР «Сервис-У». — М.: ФИЦ ИУ РАН, 2016. 437 с.
7. Зацаринный А. А., Горшенин А. К., Волович К. И., Колин К. К., Кондрашев В. А., Степанов П. В. Управление научными сервисами как основа национальной цифровой платформы «Наука и образование» // Стратегические приоритеты, 2017. № 2(14). С. 103–113.
8. Top 500. The List. June 2018. <https://www.top500.org/lists/2018/06>.
9. Zatsarinny A., Gorshenin A., Kondrashev V., Volovich K., Denisov S. Toward high performance solutions as services of research digital platform // 13th Symposium (International) "Intelligent Systems" Abstracts / Eds. K. A. Pupkov, A. I. Diveev. — SPb.: LETI, 2018. P. 59–60.
10. Зацаринный А. А., Волович К. И., Кондрашев В. А. Методологические вопросы управления научными сервисами научных и образовательных организаций Российской Федерации // Радиолокация, навигация, связь: Сб. трудов XXIII Междунар. научн.-технич. конф. РЛНС-2017. — Воронеж: Вэлборн, 2017. Т. 1. С. 7–14.
11. Информационные, управляющие и телекоммуникационные системы. Ч. 5: Исследование научно-практических вопросов создания исследовательской инфраструктуры ФИЦ ИУ РАН и ее применения для представления научных сервисов: Промежуточный отчет о НИР. — М.: ФИЦ ИУ РАН, 2017. С. 7–18.

Поступила в редакцию 20.08.18

ORGANIZATION OF CALCULATIONS IN A HYBRID HIGH-PERFORMANCE COMPUTING CLUSTER FOR PARALLEL EXECUTION OF HETEROGENEOUS TASKS

K. I. Volovich

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

Abstract: Approaches to provision of computing resources for solving applied and fundamental scientific problems are considered. The article suggests methods and algorithms for adapting and executing user applications on a hybrid computing cluster that uses specialized accelerators to perform calculations. The authors offer solutions for migrating code in high-level languages to the target platform of a hybrid computing cluster, algorithms for providing computing resources in a multitask environment, and methods for creating an adapted execution environment using virtualization technology.

Keywords: high performance computing; cluster; digital platform; virtualization; computational process control; distributed computing; parallel computing; graphics accelerator; HPC

DOI: 10.14357/08696527180410

Acknowledgments

The work was partly supported by the Russian Foundation for Basic Research (projects 18-07- 00669 and 18-29-03100).

References

1. Wu, W., H. Zhang, Zh. Li, and Y. Mao. 2011. Creating a cloud-based life science gateway. *e-Science and the Archaeological Frontier: 7th Conference (International) on eScience*. Piscataway, NJ: IEEE. 55–61.
2. Ding, F., D. Mey, S. Wienke, R. Zhang, and L. Li. 2014. A study on today’s cloud environments for HPC applications. *3rd Conference (International) on Cloud Computing and Services Science Proceedings*. Berlin, Germany: Springer. 114–127.
3. Volkov, S., and O. Sukhoroslov. 2017. Simplifying the use of clouds for scientific computing with Everest. *Procedia Comput. Sci.* 119:112–120.
4. Volovich, K. I., A. A. Zatsarinnyy, V. A. Kondrashev, and A. P. Shabanov. 2017. O nekotorykh podkhodakh k predstavleniyu nauchnykh issledovaniy kak oblachnogo servisa [Scientific research as a cloud service]. *Sistemy i Sredstva Informatiki — System and Means of Informatics* 27(1):73–84.

5. Zatsarinny, A. A., A. K. Gorshenin, V. A. Kondrashev, and K. I. Volovich. 2017. Sistema upravleniya nauchnymi servisami kak bazovyy servis tsifrovoy platformy dlya nauchnykh issledovaniy [Scientific services management system as a basic service of a digital platform for scientific research]. *Tr. 3-y nauchn.-praktich. konf.* [3rd Scientific and Practical Conference Proceedings]. Moscow: IPU RAS, NITS “Zhukovsky Institute.” 53–64.
6. FRC CSC RAS. 2016. Issledovanie voprosov upravleniya rezul'tatami nauchno-issledovatel'skoy deyatel'nosti organizatsiy, podvedomstvennykh FANO Rossii, i nauchnymi servisami seti TsKP FANO [Research of the issues of managing the results of research activities of organizations subordinate to the FANO of Russia, and the scientific services of the CKP FANO]. Otchet o NIR “Servis-U” [Report on the research work “Service-U”]. 437 p.
7. Zatsarinny, A. A., A. K. Gorshenin, K. I. Volovich, K. K. Kolin, V. A. Kondrashev, and P. V. Stepanov. 2017. Upravlenie nauchnymi servisami kak osnova natsional'noy tsifrovoy platformy “Nauka i obrazovanie” [Management of scientific services as the basis of the national digital platform “Science and Education”]. *Strategic Priorities* 2(14):103–113.
8. Top 500. The list. June 2018. Available at: <https://www.top500.org/lists/2018/06/> (accessed July 30, 2018)
9. Zatsarinny, A., A. Gorshenin, V. Kondrashev, K. Volovich, and S. Denisov. 2018. Toward high performance solutions as services of research digital platform. *13th Symposium (International) “Intelligent Systems” Abstracts*. Eds. K. A. Pupkov and A. I. Diveev. St. Petersburg: LETI. 59–60.
10. Zatsarinny, A. A., K. I. Volovich, and V. A. Kondrashev. 2017. Metodologicheskie voprosy upravleniya nauchnymi servisami nauchnykh i obrazovatel'nykh organizatsiy Rossiyskoy Federatsii [Methodological problems of management of Russian Federation scientific and educational organizations science services]. *Radiolocation, Navigation, Communication: Collection of Materials of the 23rd Scientific and Technical Conference (International)*. Voronezh: Velborn. 1:7–14.
11. FRC CSC RAS. 2017. Informatsionnye, upravlyayushchie i telekommunikatsionnye sistemy. Ch. 5. Issledovanie nauchno-prakticheskikh voprosov sozdaniya issledovatel'skoy infrastruktury FITs IU RAN i ee primeneniya dlya predstavleniya nauchnykh servisov [Information, control, and telecommunication systems. Part 5. Research of scientific and practical issues of creating the research infrastructure of FRC CSC RAS and its application for the presentation of scientific services]. Moscow. Technical Report. 7–18.

Received August 20, 2018

Contributor

Volovich Konstantin I. (b. 1970) — 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; KVolovich@frccsc.ru

ОСНОВЫ ЖИВУЧЕСТИ АВТОМАТИЗИРОВАННЫХ ОРГАНИЗАЦИЙ ИНФОРМАЦИОННОГО ОБЩЕСТВА

И. И. Быстров¹, В. Н. Веселов², К. К. Колин³

Аннотация: Рассматриваются вопросы, связанные с обеспечением живучести информационного общества, в котором широко используются информационно-телекоммуникационные системы и технологии. Ставится проблема создания методологии проектирования комплексной системы обеспечения живучести (КСОЖ) автоматизированных организаций. Рассматриваются модели и показатели обеспечения живучести информационно-телекоммуникационных систем.

Ключевые слова: информационное общество; информационно-телекоммуникационная система; живучесть автоматизированных организаций; модели и показатели обеспечения живучести

DOI: 10.14357/08696527180411

1 Актуальность проблемы

На сегодняшний день проблемы, связанные с информатизацией общества, приобрели новую актуальную значимость, так как уровень применения сетевых информационных технологий (ИТ) и информационно-коммуникационных технологий (ИКТ), который является одним из важных факторов конкурентности страны в мировой экономике, а также в политической и военной сфере, в настоящее время становится материально-технической базой цифровой экономики.

Комплексная автоматизация общества предполагает системный подход, широкое внедрение средств вычислительной техники, коммуникаций и соответствующего программного обеспечения в деятельность автоматизированных организаций, т. е. организаций, деятельность которых осуществляется на основе автоматизации функциональных процессов. В результате этого значительная часть функциональных процессов, определяющих целевое предназначение этих организаций, выполняется с использованием информационно-телекоммуникационных систем (ИТС).

При этом если ранее в случае выхода из строя автономных компьютерных систем организации сравнительно легко переходили на ручные методы управления,

¹Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, ibystrov@ipiran.ru

²Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, vveselov@ipiran.ru

³Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, kolinkk@mail.ru

то при комплексной автоматизации такой подход стал практически невозможен. Это привело к тому, что жизнедеятельность автоматизированной организации стала зависеть от состояния ИТС.

В последние годы все большее развитие и применение, особенно в крупных государственных, коммерческих и военных организациях, стали получать современные распределенные по всей структуре общества и государства информационно-телекоммуникационные системы (сети), представляющие собой интеграцию информационных и коммуникационных составляющих различных компьютерных систем (сетей) с опорой на взаимосвязанную совокупность крупных многофункциональных центров коллективной обработки информации.

Внедрение этих систем существенно повысило эффективность применения ИТ в деятельности общества, в формировании и развитии национальной цифровой экономики, обеспечении интересов личности и реализации стратегических национальных приоритетов [1]. В связи с этим приобретает особую важность и актуальность проблема обеспечения устойчивого и бесперебойного функционирования автоматизированных субъектов и объектов общества и государства на базе ИТС, их способности надежно сохранять и восстанавливать свою функциональность в условиях воздействия всего спектра угроз.

Современные ИТС представляют собой *организационно-технические структуры* с высокой концентрацией технических, программных средств и информации, которые могут быть объектами случайных или преднамеренных злоумышленных воздействий на основе использования свойства уязвимости информации, инфраструктуры систем (сетей), подсистем, процессов и оборудования.

Структурная сложность и территориальная распределенность ИТС и их компонентов обуславливают их уязвимость от множества внутренних и внешних угроз различной природы, что не позволяет превентивными мерами обеспечить их гарантированную защиту. Нарушение их функциональности приводит к возникновению высоких рисков в жизнедеятельности субъектов и объектов общества, к нарушению их работоспособности и информационной безопасности, а в некоторых случаях может приводить и к остановке деятельности критических объектов инфраструктуры общества и государства.

Наличие свойства уязвимости ИТС позволяет реализовать угрозы нарушения конфиденциальности и целостности информации, обойти превентивные меры защиты компонентов систем и сетей, нарушить их функциональные процессы и физическое состояние. Так, в 2015 г. в России было выявлено 189 уязвимостей в компонентах автоматизированных систем управления, большинство из которых являлись критическими (49%) или имели средний уровень опасности (42%).

Все это приводит к тому, что определенные субъекты (коалиции, государства, организации и личности) стремятся через инфраструктуру ИТС, компьютерные системы (сети, Интернет) осуществлять *вредоносные кибервоздействия* на жизнедеятельность общества и государства в своих интересах.

Ввиду невозможности обеспечить достаточно высокую киберзащиту общества одной из важнейших задач обеспечения национальной безопасности страны

становится придание автоматизированным организациям и всей инфраструктуре государства способности непрерывно функционировать в условиях воздействия существующих угроз и выхода из строя отдельных критичных объектов.

Согласно результатам одного из последних исследований, проведенных компанией Forester, планирование и управление непрерывностью бизнеса, а также обеспечение восстановления деятельности в непредвиденных обстоятельствах и при стихийных бедствиях определяются как два приоритетных направления деятельности организаций, которые сегодня находятся в центре внимания руководителей ведущих финансовых и других автоматизированных организаций на всех континентах [2].

При таком подходе речь должна идти о новом комплексном свойстве — *живучести* (ГОСТ 34.003-90) как отдельных автоматизированных организаций, так и живучести всего информационного общества в целом.

2 Живучесть информационного общества: содержание проблемы и методология ее решения

Под *живучестью информационного общества* будем понимать такое его состояние, при котором оно способно в условиях воздействия угроз поддерживать свою жизнедеятельность, доступность услуг, а также сохранять и восстанавливать функциональность критичных автоматизированных организаций с минимизацией ущерба информационной и национальной безопасности.

Проблему живучести информационного общества как в теоретическом, так и в практическом плане целесообразно решать комплексно. Это должно осуществляться на базе *общесистемной методологии*, объединяющей на системном уровне правовые основы Стратегии развития информационного общества в Российской Федерации на 2017–2030 гг. и Доктрины информационной безопасности РФ [3], а также последние научные достижения множества дисциплин, связанных с применением ИТ и их реализацией в виде ИТС.

Ключевым моментом методологии обеспечения живучести информатизированного общества является создание *теории и практики построения единого технологического и организационного поля*, в котором критичные системы, ИТС и автоматизированные организации общества и государства сохраняют свою способность противодействовать угрозам, осуществлять предоставление ИТ-услуг и восстанавливать свою функциональную деятельность.

В настоящее время сложилось несколько направлений решения проблемы обеспечения информационной безопасности общества с использованием средств информатизации. Представители первого направления связывают решение этой проблемы с акцентом на *техническую защиту информации* в системах, без их адаптации к угрозам технического и физического разрушения инфраструктуры, возможным рискам и последствиям их для общества и государства.

Другие рассматривают проблему влияния устойчивости систем на жизнедеятельность общества только с точки зрения *надежности технических и программных средств*.

Третьи предлагают решать эту проблему с позиции противодействия угрозам, связанным с манипулированием функциональностью системы, ресурсами, информацией, блокированием и физическим разрушением инфраструктуры и киберсреды объектов общества, органов военного и государственного управления.

По каждому из этих направлений в настоящее время выработаны определенные требования к безопасности общества и государства, но это в основном рекомендации «частного», корпоративного, технического и правового плана. Несмотря на то что эти подходы тесно связаны между собой, их раздельное применение решает проблемы информационной безопасности общества и национальной безопасности не в полной мере.

Живучесть, по мнению авторов, является *комплексным свойством* автоматизированного общества и государства и зависит от множества факторов. Среди множества свойств, которые целесообразно учитывать, исследуя жизнеспособность информатизированного общества и государства, необходимо выделять их способность существовать и сохранять свою функциональность на основе *живучести критических автоматизированных организаций, функционирующих* на базе ИТС [4].

Понятие «*живучесть автоматизированной организации*» в традиционном понимании — это функциональное свойство устойчивости сложных систем, каковыми являются современные автоматизированные субъекты информационного общества и государства.

В качестве нового определения под *живучестью автоматизированной организации* будем понимать такое ее свойство, при котором компьютерные системы (сети), интегрированные в инфраструктуру ИТС, способны поддерживать бесперебойное функционирование деятельности организации, доступность ИТ-услуг, сохраняя и восстанавливая свою функциональность с использованием специальных механизмов адаптации к условиям воздействия угроз.

Все угрозы для функционирования ИТС могут быть классифицированы по целому ряду критериев. Базельский комитет предлагает следующую классификацию источников этих угроз:

- несанкционированные действия персонала организации и некорректное выполнение своих обязанностей;
- нарушения в результате сбоя функционирования автоматизированных систем и процессов доступа к ИТ-сервисам;
- преднамеренные угрозы и действия, исходящие из внешней среды.

Наряду с традиционными угрозами, возникающими в инфраструктуре ИТС, в последние годы с распространением глобальных, национальных корпоративных ИТС крупных организаций определилась техническая среда, через которую осуществляются кибервойны, кибератаки криминальных элементов общества,

террористов, информационно-технологическое воздействие на личности и принимаемые ими решения [5]. В результате этого в зависимости от сферы применения ИТС, их уязвимости и характера угроз создаются опасности для жизнедеятельности организаций и всего общества в целом, которые по своей общей направленности в соответствии с Доктриной информационной безопасности оказывают влияние:

- на конституционные права и свободы граждан, реализуемые в информационной сфере;
- состояние инфраструктуры;
- информационные ресурсы.

История проблемы живучести деятельности автоматизированных организаций началась с момента появления электронных вычислительных машин (ЭВМ). По мере расширения сферы использования ЭВМ с появлением автоматизированных информационных систем и центров обработки данных (ЦОД) возникла проблема защиты информации в системах, расширился круг решаемых задач, а деятельность организаций стала все больше зависеть от качества и непрерывности функционирования этих систем и центров. В 1980-х гг. окончательно сформировалась проблема аварийного восстановления в случае возникновения угроз, способных нарушить функционирование систем, и в первую очередь ЦОД.

В 1990-е гг. на смену автономным ЦОД пришли персональные компьютеры и, соответственно, изменился подход к резервированию данных и аварийному восстановлению автоматизированных информационных систем. Большинство организаций отказались от централизованных мейнфреймов, заменяя их большим числом серверов и пользовательских автоматизированных мест, распределенных по всей структуре организации. Возникло множество достаточно автономных, разнородных автоматизированных систем документальной связи, различных информационных и информационно-управляющих систем.

На смену экстенсивному развитию в конце 1990-х гг. пришло интенсивное развитие в направлении создания крупномасштабных мультисервисных ИТС, интегрирующих информационные и телекоммуникационные ресурсы различных отдельных систем и сетей. Эти системы органически слились со структурой организаций и стали определять значительную долю процессов их деятельности, скорость их выполнения, и непрерывность процессов стала зависеть от бесперебойной обработки информации, ее хранения и качества сервисов и ИТ-услуг, предоставляемых ИТС. В связи с этим на смену термину «аварийное восстановление» пришел термин «непрерывность бизнеса».

В последние 10–15 лет ведущие мировые компании, в первую очередь финансового сектора, в условиях постоянно растущих угроз информационным и коммуникационным технологиям и частоты возникновения чрезвычайных ситуаций осознали степень зависимости бизнеса от состояния непрерывности функционирования и доступности сервисов ИТС. Поэтому начался процесс целенаправлен-

ного внедрения требований непрерывности деятельности организаций на базе ИТС.

В США, Великобритании, Австралии, Новой Зеландии, Японии, Сингапуре в эти годы создаются институты, которые разрабатывают стандарты и рекомендации в области «непрерывности функционирования и доступности сервисов ИТС». В результате этого понятие *непрерывности функционирования и доступности ИТ-услуг* стало базовым при решении проблем живучести автоматизированных организаций.

3 Методология обеспечения живучести автоматизированных организаций

Главная цель обеспечения живучести автоматизированной организации, функционирующей на базе ИТС, заключается в сохранении ее работоспособности в случае полного или частичного нарушения ресурсно-сервисных возможностей ИТС. Здесь важно учитывать, что в случае нарушения этих требований ИТС превращается из незаменимого средства технической поддержки штатной работы в средство, которое может парализовать всю деятельность организации. Отсюда возникает центральное требование к ИТС — *способности адаптироваться к вероятным угрозам и сохранять свою способность к функционированию*. Для достижения этих качеств ИТС должна обладать комплексным свойством живучести, которое предполагает ее способности:

- обеспечения непрерывности функционирования и доступности сервисов и ИТ-услуг;
- обеспечения комплексной безопасности инфраструктуры и информационных ресурсов;
- управление рисками, порождаемыми самой ИТС.

Основой создания комплексной системы обеспечения технической живучести автоматизированной организации должна стать *общесистемная методология*, интегрирующая в себе организационно-технический инструментарий реализации целей и задач автоматизированной организации. Общесистемная методология обеспечения живучести автоматизированной организации на базе ИТС охватывает широкий круг вопросов, в частности:

- разработку научных основ обеспечения устойчивости, применение методов системного и процессного подхода;
- создание моделей формирования и поддержки ИТ-услуг в различных условиях обстановки с ориентацией на пользователей;
- разработку принципов проектирования, комплексирования и интеграции методов и средств в инфраструктуру ИТС и ее технологических процессов;

- создание методов и способов целостного управления с основных и резервных площадок с необходимым уровнем информационной поддержки процессов принятия решений в чрезвычайных ситуациях;
- создание методики определения критически важных объектов, систем, сервисов и данных, оценки технических рисков и приоритетов восстановления компонентов ИТС, подвергшихся разрушению;
- разработку принципов построения ресурсно-сервисной системы эксплуатации для полной поддержки живучести действующих ИТС и обеспечения предоставления пользователям критически важных сервисов и ИТ-услуг в период действия чрезвычайной ситуации;
- разработку нормативно-методической базы, планов, регулирующих вопросы обеспечения живучести организации на базе ИТС и восстановления сервисов в случае возникновения угроз и чрезвычайных ситуаций.

При решении этих вопросов должны проводиться:

- анализ угроз, приводящих к нарушению живучести ИТС;
- оценка уязвимости ИТС, ее компонентов и процессов;
- анализ рисков для ИТС, порождаемых угрозами, и оценка их влияния на живучесть, надежность и помехозащищенность аппаратно-программных средств;
- определение критичности объектов, систем (сетей) и подсистем, нарушение функционирования которых оказывает влияние на способность ИТС выполнять свои функции в заданных режимах;
- анализ и оценка ущерба организации, возникающего в случае нарушения непрерывности функционирования и доступности сервисов (НФДС) ИТС.

Исходя из приведенного выше определения живучести, можно дать следующее определение общего требования к живучести функционирования любого автоматизированного субъекта общества и государства на базе ИТС.

Живучесть автоматизированного субъекта общества и государства представляет собой стратегическую, оперативную и тактическую способность планировать и осуществлять свою работу в случае возникновения негативных факторов, нарушающих его деятельность, направленную на обеспечение непрерывности и комплексной безопасности деловых операций в соответствии с установленным приемлемым уровнем рисков для ИТС.

Деятельность субъекта общества (организации) на каждом уровне следует рассматривать как комплексный сбалансированный автоматизированный процесс, охватывающий такие институциональные аспекты, как правовая база, стандарты, рыночная конъюнктура, механизмы стратегического планирования процессов обеспечения непрерывности и восстановления деятельности (ОНиВД), а также информационные и технические аспекты, включая требования к способности ИТС обеспечивать функциональные процессы организации.

На стратегическом уровне определяются принципы деятельности организации на базе ИТС, живучесть автоматизированной организации и комплекс организационно-технических мер реализации стратегии и плана действий в чрезвычайных ситуациях по защите объектов, персонала, критических функциональных процессов.

На оперативном уровне определяются концептуальные основы создания КСОЖ:

- идентификация критически важных для организации процессов и поддерживающих ресурсов ИТС;
- оценка рисков и возможного ущерба от нарушения функционирования ИТС;
- определение требований к обеспечению НФДС ИТС;
- выбор общих способов применения превентивных мер обеспечения безопасности;
- выбор методов устранения всех видов рисков, которые не удалось исключить с помощью превентивных мер, а также мер, предоставляемых на период восстановления;
- разработка требований, синхронизирующих подходы к управлению непрерывностью функционирования организации и ИТС;
- разработки методов и средств реализации требований нормативных документов и методик оценки соответствия организационных и технических решений создания КСОЖ.

На тактическом уровне определяется организационно-техническое обеспечение и поддержка процессов практического проектирования КСОЖ ИТС, решаются проблемы технических решений НФДС, эксплуатационной живучести, безопасности ресурсов и информации, а также восстановления функционирования ИТС в приемлемый для организации период времени.

Основная цель создания КСОЖ ИТС — техническая поддержка нормальной деятельности организации и восстановление нарушенной работоспособности ИТС в максимально короткие сроки с минимизацией оперативного влияния на ее процессы. Реализация мероприятий в организации на каждом уровне определяется стратегией живучести общества и государства, учитывающей необходимый баланс между экономическими, финансовыми, деловыми, эксплуатационными рисками и планированием мероприятий на случай чрезвычайных происшествий.

Организационно-технической базой восстановления функционирования ИТС являются критерии:

- быстрого и своевременного восстановления функциональности;
- высокого уровня уверенности в том, что (за счет повседневного использования и проверочного тестирования) внутренние и внешние мероприятия по обеспечению непрерывности процессов эффективны и совместимы.

При этом КСОЖ должна удовлетворять:

требованиям к времени восстановления:

- для ведущих критичных организаций — процессы восстановления и возобновления своей деятельности в течение дня с момента аварии;
- для организаций, играющих важную роль для других критически важных организаций, — возобновление деятельности в пределах нескольких часов;

требованиям к структурной устойчивости:

- иметь для восстановления после широкомасштабной катастрофы в заданное время географически разнесенные основную и резервные площадки для выполнения деловых операций и работы ЦОД;
- резервные площадки ИТС не должны зависеть от тех же инфраструктурных компонентов, что и основная, включая средства транспортировки, телекоммуникации, водоснабжения, электроснабжения и т. п.;

требованию к готовности — наличию эффективных планов обеспечения непрерывности и восстановления деятельности организации в условиях возникновения катастрофы.

4 Заключение

Развитие ИТС определяет процесс создания научно-технических и экономических основ нового типа субъектов цифрового общества и государства, у которых вся их деятельность в существенной степени зависит от состояния ИТС.

Существующие перспективные ИТС и технологии аккумулируют и обрабатывают большие объемы критической информации и могут практически контролировать и регулировать деятельность любого субъекта общества и государства. Это приводит к тому, что выживание любого субъекта, общества и государства становится зависимым от состояния ИТС и ИТ, а сами системы стали объектом воздействия различного рода угроз и влиять на жизнеспособность общества и государства.

В условиях существования информатизированного общества и государства в среде современных реальных и потенциальных угроз ИТС, когда методы и средства защиты будут отставать от процессов появления новых угроз, а последствия от нарушения функциональности ИТС будут постоянно возрастать, *проблема обеспечения их живучести* приобретает особую важность и актуальность для национальной безопасности Российской Федерации.

В настоящее время эта проблема еще не нашла должного научного исследования как multidisciplinary комплексная проблема. В составе этой проблемы необходимо в первую очередь выделять ее технологические аспекты:

- вопросы категорирования критически важных объектов, систем и сервисов;
- вопросы обеспечения комплексной безопасности ИТС, резервирования компонентов системы, требований к восстановлению и эксплуатационной поддержке живучести, к устойчивости функциональных процессов;
- вопросы создания комплексной системы обеспечения живучести ИТС.

Однако в дальнейшем необходимо будет развернуть комплексные междисциплинарные исследования этой проблемы, которые будут включать и другие, в том числе гуманитарные, аспекты этой проблемы. В их числе:

- вопросы управления непрерывностью и операционными рисками, порождаемыми ИТС, для жизнедеятельности общества и государства;
- вопросы защиты материальных и финансовых ресурсов, репутации на рынке, защиты сотрудников, инфраструктуры субъектов общества и государства от нарушений и рисков, порождаемых киберугрозами.

В условиях развития цифровой экономики и становления глобального информационного общества зависимость безопасности жизнедеятельности общества от качества функционирования ИТС и уровня живучести автоматизированных организаций будет, безусловно, возрастать. Поэтому научные исследования и прикладные разработки, направленные на решение этой актуальной проблемы, должны стать важным направлением мероприятий по укреплению национальной безопасности нашей страны [6].

Литература

1. *Коллин К. К.* Информатизация общества и глобализация. — Красноярск: СФУ, 2011. 52 с.
2. Доклад компании Forrester Report Sun Gard от 02.09.2010.
3. Доктрина информационной безопасности Российской Федерации: Утверждена Указом Президента РФ от 5 декабря 2016 г. № 646.
4. *Быстров И. И.* Живучесть автоматизированных организаций. — М.: Майор, 2016. 506 с.
5. *Коллин К. К.* Технологическое общество: глобальные тенденции, вызовы и угрозы // Стратегические приоритеты, 2017. № 1. С. 4–15.
6. *Соколов И. А., Коллин К. К.* Развитие информационного общества в России и актуальные проблемы информационной безопасности // Информационное общество, 2009. № 4-5. С. 98–107.

Поступила в редакцию 05.06.18

INFORMATION SOCIETY AUTOMATED ORGANIZATIONS SURVIVABILITY BASICS

I. I. Bystrov, V. N. Veselov, and K. K. Kolin

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

Abstract: The paper deals with the issues related to ensuring survivability of the information society in which information and telecommunication systems and technologies are widely used. The problem of creating a methodology for designing an integrated system for ensuring survivability of automated organizations is raised. Models and indicators of ensuring survivability of information and telecommunication systems are considered.

Keywords: information society; information-telecommunication system; automated organizations survivability; models and indicators of ensuring survivability

DOI: 10.14357/08696527180411

References

1. Kolin, K. K. 2011. *Informatizatsiya obshchestva i globalizatsiya* [Society informatization and globalization]. Krasnoyarsk: SFU Publ. 52 p.
2. Doklad kompanii Forrester [Company Forrester Report]. 02.09.2010. Sun Gard.
3. Doktrina informatsionnoy bezopasnosti Rossiyskoy Federatsii [Doctrine of information security of the Russian Federation]. 2016. Utverzhdena Ukazom Prezidenta RF ot 5 dekabrya 2016 g. No. 646 [Approved by Decree of the President of RF of December 5, 2016, No. 646].
4. Bystrov, I. I. 2016. *Zhivuchest' avtomatizirovannykh organizatsiy* [Automated organization survivability]. Moscow: Mayor Publ. 506 p.
5. Kolin, K. K. 2017. Tekhnologicheskoe obshchestvo: global'nye tendentsii, vyzovy i ugrozy [Technological society: Global tendencies, calls, and threats]. *Strategic Priorities* 13(1):4–15.
6. Sokolov, I. A., and K. K. Kolin. 2009. Razvitie informatsionnogo obshchestva v Rossii i aktual'nye problemy informatsionnoy bezopasnosti [Information society development in Russia and information security actual problems]. *Informatsionnoe obshchestvo* [Information Society] 21(4-5):98–107.

Received June 5, 2018

Contributors

Bystrov Igor I. (b. 1931) — Doctor of Science in technology, professor, principal scientist, Institute of Informatics Problem, Federal Research Center “Computer

Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333; ibystrov@ipiran.ru

Veselov Vitaliy N. (b. 1940) — Candidate of Science (PhD) in technology, leading engineer, Institute of Informatics Problem of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; vveselov@ipiran.ru

Kolin Konstantin K. (b. 1935) — Doctor of Science in technology, professor, Honored Scientist of the Russia Federation, principal scientist, Institute of Informatics Problem, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; kolinkk@mail.ru

РАЗРАБОТКА ПРИЛОЖЕНИЯ ДЛЯ ПОДГОТОВКИ РАСЧЕТНЫХ СЕТОК С ГРАДУИРУЮЩИМИ И ИЗОГНУТЫМИ КРАЯМИ ДЛЯ ПРОГРАММНОЙ СРЕДЫ OpenFOAM

Д. И. Читалов¹, С. Т. Калашиников²

Аннотация: Статья посвящена проблеме упрощения подготовки расчетных сеток (РС) с градуирующими и изогнутыми краями для программной среды (ПС) OpenFOAM. Авторами рассмотрены имеющиеся на текущий момент коммерческие и свободно распространяемые программные средства для решения обозначенной проблемы. Определены недостатки этих программных средств, актуальность проблемы и необходимость создания собственного приложения. Представлен список стоящих перед авторами задач и описание объекта исследования. Сформулированы вызовы, стоящие перед авторами в процессе реализации разработки, и предложены пути их решения. Представлены результаты применения программы на примере одной из учебных задач, входящих в ПС OpenFOAM. Сформулированы итоги проведенной работы, определены ее практическая значимость и перспективы последующих исследований в данном направлении.

Ключевые слова: графический пользовательский интерфейс; расчетные сетки с градуирующими и изогнутыми краями; OpenFOAM; язык программирования Python; библиотека PyQt4; открытое программное обеспечение

DOI: 10.14357/08696527180412

1 Введение

В настоящее время при проведении экспериментов по численному решению задач механики сплошных сред (МСС), прежде всего таких ее областей, как гидро-, газодинамика и механика деформируемого твердого тела, широко применяются программные продукты математического моделирования. Среди таких пакетов программ популярностью пользуются коммерческие решения с закрытым исходным кодом (ANSYS, ТЕСИС, ESI GROUP) и свободно распространяемые комплексы, в частности OpenFOAM [1].

Несмотря на то что ПС OpenFOAM не требует приобретения лицензии и успешно используется при постановке численных экспериментов, этот программный комплекс не лишен недостатков. Главным из них является необходимость работы с программой через терминал, что требует от пользователя

¹Южно-Уральский научный центр, г. Миасс, Челябинская обл., cdi9@yandex.ru

²Южно-Уральский научный центр, г. Миасс, Челябинская обл., src@makeyev.ru

Сравнительная характеристика популярных графических оболочек для работы с ПК OpenFOAM

Название	Лицензия	Техническая поддержка	Русскоязычная документация
Salome [3]	ПО с открытым кодом	Не предусмотрена	Отсутствует
Helyx-OS [4]	ПО с открытым кодом	Предусмотрена, платная	Отсутствует
Visual-CFD [5]	Бесплатная версия на 30 дней	Предусмотрена, платная	Отсутствует

знания структуры и допустимого содержимого служебных файлов, входящих в расчетный случай, а также знания консольных команд.

Настоящая статья является продолжением работы [2], в которой авторами предложена собственная графическая оболочка для проведения этапов численного моделирования задач МСС на базе ПК OpenFOAM, в том числе подготовки расчетного случая, импорта готовой РС, запуска процесса решения задачи МСС и визуализации результатов. В настоящей статье рассматривается проблема упрощения этапа подготовки РС, в частности РС с градуирующими и изогнутыми краями, и в рамках решения поставленной проблемы — разработка приложения с графическим пользовательским интерфейсом для подготовки РС данного типа.

Над проблемой упрощения работы пользователя при проведении математических исследований посредством ПК OpenFOAM работают программисты всего мира и даже целые компании, запустившие в связи с ростом интереса к ПК OpenFOAM со стороны разработчиков и инженеров собственные проекты по разработке графических пользовательских оболочек для работы с данным программным комплексом.

В таблице представлена сравнительная характеристика ряда таких оболочек, получивших наибольшую популярность у мирового сообщества исследователей и имеющих в своем составе модули подготовки РС с градуирующими и изогнутыми краями.

Для представленных программных комплексов возможно использование без приобретения лицензии. Вместе с тем их практическое применение российскими пользователями сопряжено с некоторыми сложностями, прежде всего с отсутствием официальной русскоязычной документации и с необходимостью приобретения услуг по технической поддержке.

Резюмируя вышесказанное, можно сделать вывод о том, что освоение и использование имеющихся на текущий момент графических программных решений для работы с ПК OpenFOAM на отечественных промышленных предприятиях связано с существенными временными и финансовыми затратами.

С перечисленными в имеющихся графических оболочках недостатками столкнулись специалисты предприятия АО «ГРЦ Макеева» при постановке численных

экспериментов по моделированию задач гидро- и газодинамики, в частности при подготовке РС с градуирующими и изогнутыми краями.

Для решения проблемы перед авторами настоящей статьи была поставлена задача создания собственного программного решения, которое бы упростило процесс подготовки РС данного типа. При этом рассматриваемое приложение для подготовки РС может представлять интерес и иметь практическую пользу и для специалистов других предприятий, работающих с ПС OpenFOAM и нацеленных на экономию рабочего времени при подготовке РС.

Объектом исследования выступает процесс подготовки параметрических сеток с градуирующими и изогнутыми краями для проведения математических исследований на базе ПС OpenFOAM. Авторами настоящей статьи проанализирована соответствующая утилита ПС OpenFOAM, отвечающая за запуск процесса генерации РС, а также необходимая для этого структура расчетного случая. На основе официальной документации [6] по ПС OpenFOAM изучена требуемая структура служебного файла, содержащего исходные данные для генерации РС, а также необходимые поля данных и допустимые типы данных.

Цель настоящей работы заключается в создании программного приложения с графическим интерфейсом, позволяющего осуществлять централизованную подготовку РС с градуирующими и изогнутыми краями для решения задач МСС с помощью ПС OpenFOAM. Согласно поставленной цели и на основе анализа объекта исследования перед авторами сформулированы следующие задачи:

1. Создать основное окно программы с русско- и англоязычной версиями интерфейса.
2. Реализовать в основном окне 4 интерфейсных блока: панель управления, блок подготовки исходных данных для сетки, блок вывода результатов и панель вывода служебных сообщений.
3. На панели управления создать кнопку, отвечающую за открытие модального окна выбора директории расчетного случая, две кнопки, отвечающие за запуск утилиты генерации и визуализации РС, и кнопку, отвечающую за открытие окна выбора языка интерфейса.
4. В блоке подготовки исходных данных для сетки создать окно с элементами управления пользовательским интерфейсом для указания исходных данных.
5. Для элементов управления реализовать возможность ввода только данных допустимого типа, реализовать систему всплывающих подсказок.
6. В блоке вывода результата создать окно для вывода содержимого сгенерированного файла с исходными данными РС и для вывода результатов генерации и визуализации РС.
7. На панели вывода служебных сообщений реализовать систему оповещения пользователя об успешных результатах выполняемых им действий и о результатах с ошибками.

8. Реализовать возможность создания для отдельного расчетного случая нескольких параметрических сеток и возможность редактирования любой из них.
9. Обеспечить визуализацию РС посредством программы научной трехмерной визуализации ParaView [7].

2 Анализ объекта исследования

Утилита «blockMesh» генерирует РС с градуирующими и изогнутыми краями, считывая исходные данные (геометрию и свойства сетки) из файла «blockMeshDict». Данный файл пользователь создает в директории «system» расчетного случая, вручную заполняет исходными данными и посредством терминала запускает команду генерации сетки.

Изучив официальную документацию по ПС OpenFOAM и учебные задачи [8], авторы статьи выделили следующие особенности работы утилиты «blockMesh», необходимые для разработки модуля подготовки параметрических сеток с градуирующими и изогнутыми краями:

- (1) утилита разбивает геометрию расчетной области на один или совокупность шестигранников в пространстве;
- (2) ребра шестигранников задаются в виде прямых линий, дуг или сплайновых кривых;
- (3) РС генерируется в зависимости от числа ячеек в каждом направлении шестигранника;
- (4) каждый шестигранник задается в местной правой системе координат (x_1, x_2, x_3) и определяется списком из 8 точек, имеющих нумерацию с 0;
- (5) файл «blockMeshDict» имеет 6 блоков исходных данных, каждый из которых определяется соответствующим ключевым словом:
 - «convertToMeters». Блок содержит единственное значение и задает коэффициент масштабирования координат вершин;
 - «vertices». Набор списков координат для указания вершин шестигранников из расчетной области;
 - «edges». Блок представляет собой список ребер шестигранников. Каждое ребро по умолчанию задается в виде прямой линии. Для выбора изогнутой формы предлагается указать одно из следующих ключевых слов: «spline» (сплайновая кривая), «polyLine» (набор линий), «BSpline» (B-сплайновая кривая) и координаты набора точек интерполяции;
 - «blocks». Блок определяет шестигранники из расчетной области. Каждому шестиграннику (ключевое слово «hex») соответствуют 8 меток

вершин, число ячеек для каждого из трех направлений (x_1, x_2, x_3) и степень расширения ячеек применительно к каждому направлению, которая позволяет «разбить» сетку в каждом из направлений. В версиях ПС OpenFOAM начиная с 2.4 предусмотрена опция мультиградуировки, отвечающая за разбиение шестигранника в одном из трех направлений на совокупность секторов и определение параметров градуировки в каждом секторе;

- «boundary». Блок определяет совокупность областей (патчей), образующих границу РС. Для каждого из патчей пользователь задает имя, тип и набор граней, формирующих патч;
- «mergePatchPairs». Блок содержит набор пар патчей, грани которых необходимо объединить. Данный блок содержит параметры только в том случае, если РС создается с использованием нескольких шестигранников и их соединение осуществляется посредством слияния граней.

На основе проведенного анализа объекта исследования авторами настоящей работы предложен подход, согласно которому работа пользователя с программой начинается с выбора директории сохранения новой РС или директории одной из существующих РС и подготовки базового описания сетки, на основе которого формируется общая структура итогового файла параметров РС «blockMeshDict».

Базовое описание предусматривает определение коэффициента масштабирования, количества используемых вершин шестигранников и числа самих шестигранников, указание наличия изогнутых ребер и определение их числа, указание наличия патчей для границы сетки и определение их числа, указание наличия слияний патчей и определение их числа. После сохранения базового описания сетки на его основе автоматически формируются формы для каждого из блоков файла «blockMeshDict» с соответствующим набором элементов управления.

3 Подготовка к разработке

Сформулированный список задач ставит перед авторами настоящей статьи следующие вызовы: выбор оптимального языка программирования и среды разработки, а также ряд вызовов алгоритмического, структурного и логического характера.

3.1 Выбор инструментария разработки

Современные настольные программные продукты разрабатываются с помощью широкого спектра высокоуровневых языков программирования, например C/C++, Java, Python. Язык программирования Python [9] считается одним из наиболее популярных инструментов реализации логики работы программных решений с графическим интерфейсом благодаря наличию встроенных структур

данных, словарей, кортежей, поддержке большого числа библиотек и простоте освоения.

Для проектирования графического интерфейса приложений, функционирующих на основе языка программирования Python, применяются различные графические библиотеки, среди которых интересом со стороны разработчиков пользуется расширение PyQt4 [10], позволяющее создавать оконные программные решения с привлекательным дизайном.

Описываемое в настоящей статье приложение для подготовки РС с градуирующими и изогнутыми краями («mesh_generator») разработано на основе языка программирования Python 3.4 и расширения PyQt4 в среде разработки IDLE. В качестве дополнительного программного обеспечения (ПО) для использования программы на компьютере под управлением операционной системы (ОС) Linux должны быть установлены ПС OpenFOAM и пакет трехмерной научной визуализации ParaView.

3.2 Анализ вызовов алгоритмического характера

К вызовам алгоритмического характера относится обеспечение возможности изменения параметров существующих РС. Для реализации данной возможности принято решение использовать модуль «pickle» языка программирования Python для «консервации» параметров определенной РС в набор служебных файлов.

В модуле «pickle» реализована мощная система сериализации [11] (запаковки) и десериализации (распаковки) объектов, созданных с помощью языка Python. Данный модуль при сериализации преобразует необходимый объект в последовательность байтов, а при десериализации — обратно в Python-объект.

При необходимости изменить параметры уже созданной РС пользователю требуется открыть директорию РС с файлами, созданными модулем «pickle», и изменить параметры сетки посредством набора форм. После этого выполнить сохранение содержимое файла «blockMeshDict» и запустить повторную генерацию и визуализацию РС. На рис. 1 представлена блок-схема, отражающая алгоритм работы пользователя с приложением «mesh_generator».

3.3 Анализ вызовов структурного характера

В результате анализа вызовов структурного характера было определено, что для описываемого программного приложения необходимо реализовать форму ввода пользователем исходных параметров РС. Ввиду сложной структуры файла «blockMeshDict» создание общей формы для всех блоков данных файла требует написания значительного объема программного кода, что может затруднить последующее изменение внешнего вида формы и логики ее работы.

В рамках решения данной проблемы принято решение для подготовки файла «blockMeshDict» создать окно с набором форм, соответствующим блокам

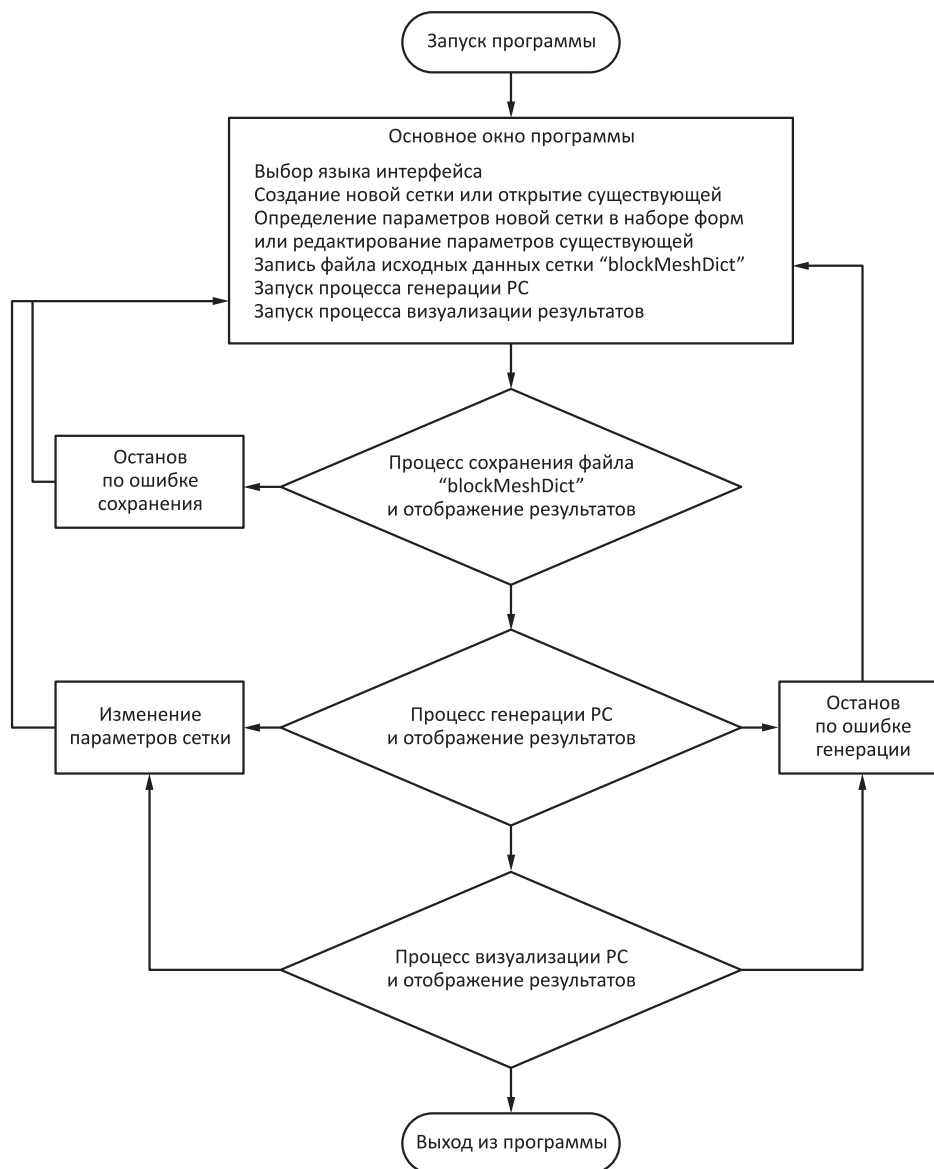


Рис. 1 Алгоритм работы пользователя с программой «mesh-generator»

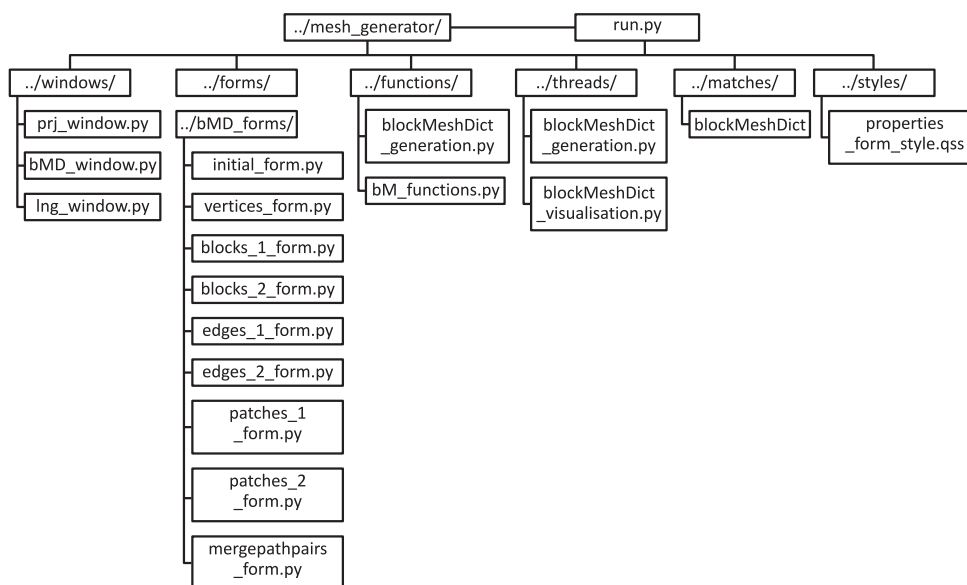


Рис. 2 Файловая структура программы «mesh_generator»

параметров РС, и отделить внешнее представление формы от логики ее работы. Согласно предложенному подходу каждой из форм должен соответствовать отдельный файл, в котором находится Python-класс, описывающий структуру элементов управления формы. Класс каждой формы импортируется в общий программный файл, отвечающий за отображение окна набора форм и сериализацию каждого соответствующего блока параметров.

На рис. 2 приведена диаграмма файловой структуры описываемого программного приложения. В директории программы находятся исполняемый файл «run.py» и 6 поддиректорий, в которых хранятся служебные файлы программы.

3.4 Анализ вызовов логического характера

Главным вызовом логического характера для авторов стала реализация возможности программного выполнения консольных команд генерации («blockMesh») и визуализации («paraFoam») РС. Для решения данной проблемы авторами предложен подход использования bash-скриптинга [12], при котором последовательность действий, выполняемых пользователем в командном интерпретаторе, записывается в файл, называемый сценарием (скриптом) командного интерпретатора (bash-script). В листинге 1 представлен bash-скрипт, автоматизирующий запуск утилиты «blockMesh».

Листинг 1. Файл сценария генерации РС на основе утилиты «blockMesh»

```
#!/bin/sh
./opt/openfoam4/etc/bashrc
blockMesh
exit
```

Как правило, отдельной программе или сценарию соответствует отдельный процесс, поэтому принято решение осуществлять процессы генерации и визуализации РС в дочерних процессах, вынесенных в потоки, параллельные основному потоку. В листинге 2 приведен код запуска bash-сценария с утилитой ПС OpenFOAM как стороннего процесса.

Листинг 2. Код запуска bash-сценария в отдельном процессе

```
msh_run_subprocess = subprocess.Popen(
[<<bash << + 'путь до скрипта'],
cwd = 'путь до расчетного случая', shell = True,
stdout='путь до файла с результатами',
stderr='путь до файла с ошибками'
)
```

Резюмируя проанализированные вызовы алгоритмического, структурного и логического характера и описанные пути реализации соответствующих задач, в рамках обоснования научной новизны разработки могут быть предложены следующие подходы:

1. Файл «blockMeshDict» с параметрами РС создается на основе данных из набора форм, соответствующих блокам параметров сетки в файле. Первая форма предназначена для инициализации РС, т. е. для формирования ее базового описания. После сохранения указанных в первой форме параметров происходит автоматическое создание других форм набора и формирование на основе этих параметров элементов управления каждой формы.
2. Каждой форме набора соответствует отдельный Python-класс, который хранится в отдельном файле, что упрощает изменение содержимого и логики работы форм в дальнейшем. Классы импортируются в общий файл, отвечающий за формирование окна блока подготовки исходных данных с набором форм и сериализацию («консервирование») данных.
3. Возможность изменения существующих сеток реализована благодаря использованию модуля «консервации» данных «pickle» языка программирования Python. «Упакованные» объекты каждого блока параметров хранятся в отдельном файле. Набор таких файлов хранится в директории, название которой совпадает с названием сетки. Пользователю для изменения уже созданной сетки необходимо выбрать эту директорию, соответствующую определенной РС, а после изменения параметров сетки выполнить сохранение файла «blockMeshDict», генерацию сетки и ее визуализацию. Таким

образом, выбранная сетка становится текущей для данного расчетного случая. Благодаря такому подходу пользователь может создавать любое число параметрических сеток для одного расчетного случая.

4. Программный запуск утилит генерации («blockMesh») и визуализации («paraFoam») РС осуществляется посредством запуска bash-сценариев с соответствующими командами как сторонних процессов из директории расчетного случая средствами модуля «subprocess» языка программирования Python.

4 Разработка

Приложение для подготовки параметрических сеток с градуирующими и изогнутыми краями («mesh_generator»), разработанное авторами для ПС OpenFOAM, является ПО со свободно распространяемым исходным кодом.

Дистрибутив программы размещен авторами на веб-сервисе для хостинга IT-проектов GitHub [13]. В дистрибутиве помимо файлов с исходным кодом приложения находятся текстовые файлы с инструкциями для русскоязычной и англоязычной версий программы.

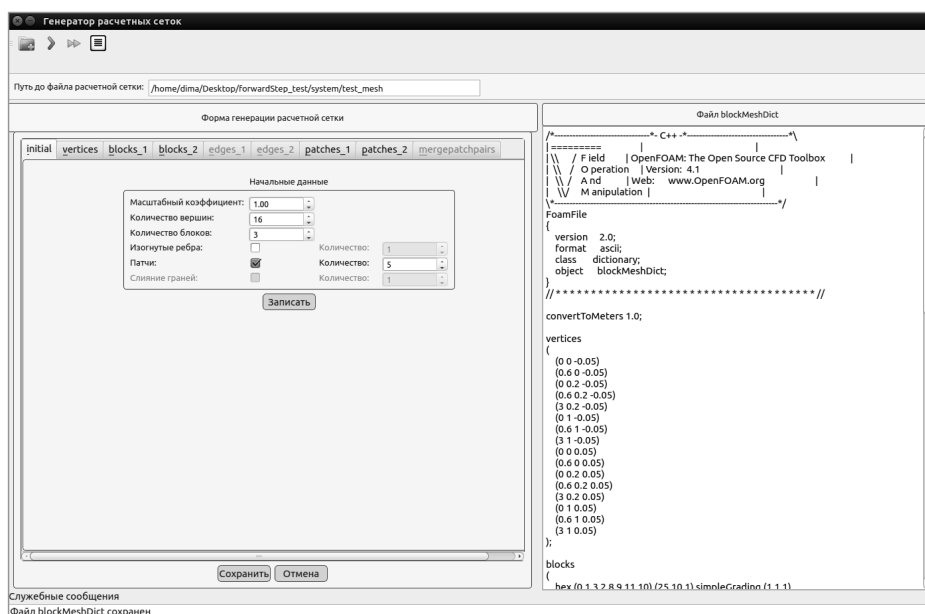


Рис. 3 Основное окно программы «mesh_generator» после сохранения файла параметров сетки «blockMeshDict»

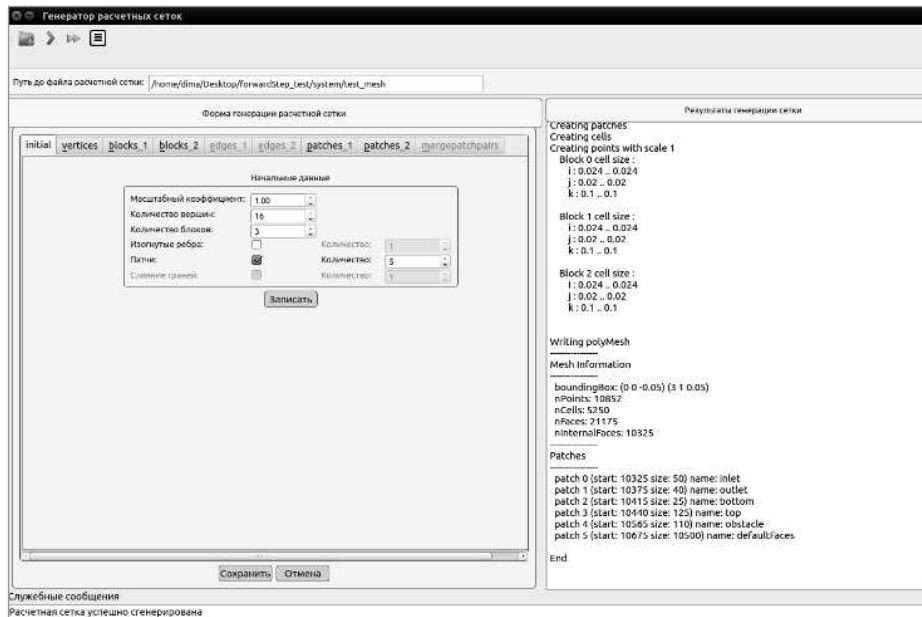


Рис. 4 Основное окно программы «mesh-generator» с результатами генерации РС

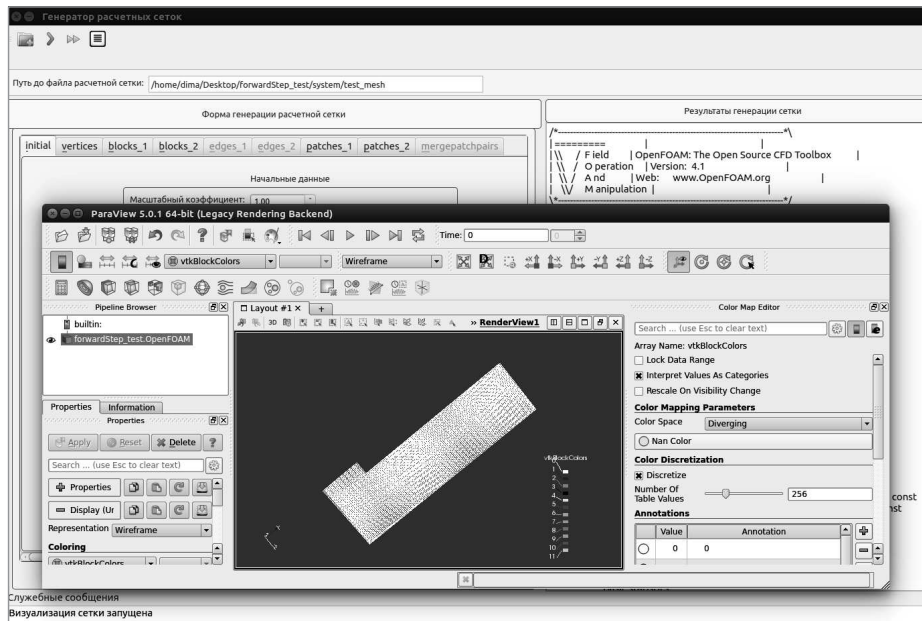


Рис. 5 Основное окно программы «mesh-generator» с визуализацией РС

Работа созданного приложения протестирована авторами на одном из учебных примеров, входящих в стандартный дистрибутив ПС OpenFOAM 4.0 (.../openfoam4/tutorials/compressible/rhoCentralFoam/forwardStep). В рамках решения задачи МСС, соответствующей данному примеру, перед пользователем ставится задача подготовки параметрической сетки посредством утилиты «blockMesh». На рис. 3 представлено основное окно программы для данного примера после сохранения файла «blockMeshDict» с параметрами РС. На рис. 4 в основном окне программы в блоке вывода результатов отображены результаты генерации РС. На рис. 5 представлено основное окно программы для описываемого примера РС с результатами запуска процесса визуализации сетки через пакет ParaView.

5 Заключение

Настоящая статья посвящена исследованию особенностей подготовки РС с градуирующими и изогнутыми краями при проведении экспериментов по численному моделированию задач МСС в ПС OpenFOAM. Авторами работы создано программное приложение «mesh_generator», позволяющее автоматизировать процесс подготовки РС. Приложение размещено в свободном доступе. Пользователи имеют возможность загрузить его и использовать для построения собственных параметрических сеток.

По результатам тестирования программы специалисты АО «ГРЦ им. Макеева» отметили значительную экономию рабочего времени при подготовке данного типа РС по сравнению с традиционным подходом без использования приложений с графическим интерфейсом. Использование приложения «mesh_generator» позволяет избавить пользователя от рутинных и затратных по времени операций и использовать сэкономленное время для решения других рабочих задач.

Дальнейшие перспективы работы авторы видят в исследовании других типов параметрических сеток для ПС OpenFOAM и утилит их генерации и последующую разработку программных решений, позволяющих автоматизировать подготовку таких сеток.

Литература

1. OpenFOAM. The open source CFD toolbox. <https://www.openfoam.com>.
2. Читалов Д. И., Меркулов Е. С., Калашиников С. Т. Разработка графического интерфейса пользователя для программного комплекса OpenFOAM // Программная инженерия, 2016. Вып. 12. С. 568–574. doi: 10.17587/prin.7.568-574.
3. Salome. The Open Source Integration Platform for Numerical Simulation, 2017. <http://www.salome-platform.org>.
4. Helyx-OS. The market leading open-source GUI for OpenFOAM, 2017. <http://engys.com/products/helyx-os>.
5. Visual-CFD for OpenFOAM, 2018. <https://www.esi-group.com/software-solutions/virtual-environment/cfd-multiphysics/visual-cfd-openfoam>.

6. OpenFOAM. User Guide, 2018. <http://foam.sourceforge.net/docs/Guides-a4/OpenFOAMUserGuide-A4.pdf>.
7. ParaView, 2017. <https://www.paraview.org>.
8. OpenFOAM. Tutorial Guide, 2017. <https://www.openfoam.com/documentation/tutorial-guide/index.php>.
9. Прохоренок Н. А. Python 3 и PyQt. Разработка приложений. — СПб.: БХВ-Петербург, 2012. 704 с.
10. PyQt4 Reference Guide, 2018. <http://pyqt.sourceforge.net/Docs/PyQt4>.
11. Сериализация объектов Python, 2018. <http://python-3.ru/page/module-pickle-python>.
12. Искусство написания Bash-скриптов, 2018. https://www.opennet.ru/docs/RUS/bash_scripting_guide.
13. Mesh-generator, 2018. https://github.com/DmitryChitalov/mesh_generator.

Поступила в редакцию 19.03.18

DEVELOPMENT OF AN APPLICATION FOR PREPARATION OF COMPUTATIONAL MESHES WITH GRADUATING AND CURVED EDGES FOR THE OpenFOAM SOFTWARE

D. I. Chitalov and S. T. Kalashnikov

Federal State Budget Scientific Institution “South Ural Scientific Center” of the Russian Academy of Sciences, Ilmen reserve, Miass, Chelyabinsk Region 456317, Russian Federation

Abstract: This article is devoted to the problem of simplifying the preparation of computational meshes (CM) with graduating and curved edges for the OpenFOAM software environment (SE). The currently available commercial and freely distributed software used to solve the indicated problem is considered. The limitations of these software tools, the urgency of the problem, and the need to create own application are identified. The list of tasks facing the authors and the description of the research object are presented. The challenges faced by the authors in the process of implementing the application are formulated and the ways of their solution are suggested. The results of application of the program by the example of one of the training tasks included in the OpenFOAM SE are presented. The results of the work, its practical significance, and prospects for subsequent research in this direction are determined.

Keywords: graphical user interface; computational meshes with graduating and curved edges; Python programming language; PyQt4 library; open source software

DOI: 10.14357/08696527180412

References

1. OpenFOAM. The open source CFD toolbox. Available at: <https://www.openfoam.com/> (accessed September 1, 2017).
2. Chitalov, D. I., E. S. Merkulov, and S. T. Kalashnikov. 2016. Razrabotka graficheskogo interfeysa pol'zovatelya dlya programmogo kompleksa OpenFOAM [Development of a graphical user interface for the OpenFOAM software package]. *Programmnyaya inzheneriya* [Software Engineering] 12:568–574. doi: 10.17587/prin.7.568-574.
3. Salome. The Open Source Integration Platform for Numerical Simulation. 2017. Available at: <http://www.salome-platform.org> (accessed September 20, 2017).
4. Helyx-OS. The market leading open-source GUI for OpenFOAM. 2017. Available at: <http://engys.com/products/helyx-os> (accessed September 20, 2017).
5. Visual-CFD for OpenFOAM. 2018. Available at: <https://www.esi-group.com/software-solutions/virtual-environment/cfd-multiphysics/visual-cfd-openfoam> (accessed September 20, 2018).
6. OpenFOAM. 2018. User Guide. Available at: <http://foam.sourceforge.net/docs/Guides-a4/OpenFOAMUserGuide-A4.pdf> (accessed October 5, 2017).
7. ParaView. 2017. Available at: <https://www.paraview.org/> (accessed October 5, 2017).
8. OpenFOAM. 2017. Tutorial Guide. Available at: <https://www.openfoam.com/documentation/tutorial-guide/index.php> (accessed October 5, 2017).
9. Prohorenok, N. A. 2012. *Python 3 i PyQt. Razrabotka prilozheniy* [Python 3 and PyQt. Application development]. St. Petersburg: BHV-Petersburg. 704 p.
10. PyQt4 Reference Guide. 2018. Available at: <http://pyqt.sourceforge.net/Docs/PyQt4/> (accessed January 10, 2018).
11. Serializatsiya ob"ektov Python [Serialization of Python objects]. 2018. Available at: <http://python-3.ru/page/module-pickle-python> (accessed January 1, 2018).
12. Iskusstvo napisaniya Bash-skriptov [The art of writing Bash scripts]. 2018. Available at: https://www.opennet.ru/docs/RUS/bash_scripting_guide/ (accessed February 1, 2018).
13. Mesh_generator. 2018. Available at: https://github.com/DmitryChitalov/mesh_generator (accessed March 15, 2018).

Received March 19, 2018

Contributors

Chitalov Dmitry I. (b. 1989) — junior scientist, South Ural Scientific Center of the Russian Academy of Sciences, Ilmen reserve, Miass, Chelyabinsk Region 456317, Russian Federation; cdi9@yandex.ru

Kalashnikov Sergey T. (b. 1956) — Candidate of Science (PhD) in technology, Head of Department, South Ural Scientific Center of the Russian Academy of Sciences, Ilmen reserve, Miass, Chelyabinsk Region 456317, Russian Federation; src@makeyev.ru

СИТУАЦИОННОЕ ПЛАНИРОВАНИЕ ПРОИЗВОДСТВА В СЕТЕВОЙ M2M-СИСТЕМЕ

А. В. Ильин¹, В. Д. Ильин²

Аннотация: Предложена методология ситуационного планирования производства (электроэнергии, воды и др.) в сетевой M2M (machine-to-machine) системе. Методология включает методы построения модели объекта по данным измерений, формирования портретов ситуаций (целевой, достигнутой и стартовой) и интерактивного планирования производства продукции. Модель объекта описывает экономические характеристики узлов и распределение потоков продукта, перемещаемого по дугам сети, в зависимости от количества продукта, отправляемого из производящих узлов и доставляемого в узлы потребления. На каждом шаге планирования постановка задачи зависит от результатов анализа портретов ситуаций. Основной составляющей вычислительного ядра служит метод целевого перемещения решения. Методология ориентирована на реализацию в виде интернет-сервисов и рассчитана на применение в сетевых M2M-системах.

Ключевые слова: сетевая M2M-система; ситуационное планирование производства; система обязательных и ориентирующих требований; метод целевого перемещения решения; интернет-сервисы

DOI: 10.14357/08696527180413

1 Введение

Задача планирования производства издавна входит в число научных проблем с неубывающей актуальностью [1–7]. Методы планирования производства в сетевых системах имеют особое значение для управления электроэнергетическими системами [8]. Сетевые M2M-системы все шире применяются в электроэнергетике, водоснабжении и других областях жизнедеятельности, информатизация которых реализуется в соответствии с Программой цифровой экономики России [9]. Исходим из того, что в современных сетевых M2M-системах [10–12] цифровые двойники [13], специализирующиеся на определенных задачах, могут быть использованы для сбора данных, формирования и анализа портретов ситуаций [14, 15] и планирования производства продукции. В предложенной методологии самообучающиеся цифровые двойники, размещенные на серверах узлов сетевой M2M-системы, предназначены для моделирования объекта (включая распределение потоков по дугам [16], экономические характеристики

¹Государственный научно-исследовательский институт авиационных систем, ilyin@res-plan.com

²Вычислительный центр им. А. А. Дородницына Федерального исследовательского центра «Информатика и управление» Российской академии наук, vdilyin@yandex.ru

производящих узлов и др.), наблюдения за его состоянием, формирования и анализа портретов ситуаций, формирования систем обязательных и ориентирующих требований [17, 18], оценки решений.

Каково состояние объекта и человеко-машинной среды решения задач, какие технологические и экономические требования целесообразно включить в систему обязательных и ориентирующих требований при поиске эффективного плана? Ответы на подобные вопросы могут быть получены путем интерпретации запросов на модели объекта ситуационного управления, представленной портретами ситуаций.

Данная статья является тематически связанным продолжением статьи, посвященной построению модели распределения потоков в однопродуктовой сетевой системе [16].

Запись формул и выделение фрагментов текста

Для записи формул, выделения определений и замечаний используются средства языка TSM-комплекса (TSM: Textual Symbolic Modeling) для описания текстовых моделей [19].

В данной статье используются следующие средства выделения фрагментов текста:

- ⟨фрагмент описания⟩ □ \approx утверждение (определение, аксиома и др.) (здесь и далее символ \approx заменяет слово «означает»);
- ◇ ⟨фрагмент описания⟩ ◇ \approx замечание.

Курсивом выделены первые вхождения названий понятий и фрагменты описания, к которым авторы хотят привлечь внимание.

2 Ситуации и их портреты

□ *Целевой* называем ситуацию, которую планируется создать в результате управляющего воздействия на объект ситуационного управления; *стартовой* — применительно к которой проектируется управляющее воздействие; *достигнутой* — полученную в результате управляющего воздействия. □

□ *Портрет ситуации* — описание, включающее данные о состоянии объекта, располагаемых вариантах управляющих воздействий и ресурсного обеспечения их реализации [14, 15]. □

Комплекс формирования и анализа портретов ситуаций работает в соответствии с *системой правил*: $\langle R, S \rangle$, где R — множество правил, $S \subseteq R \times R$ — бинарное отношение *специализации* на R . *Правило* имеет вид либо $D_1 \Rightarrow D_2$, либо $D \Rightarrow I$ (здесь D , D_1 и D_2 — произвольные описания; I — любая инструкция).

Если для любых двух правил r_1 и r_2 указано, что r_2 является *специализацией* r_1 , то r_1 является *обобщением* r_2 . Каждое *правило-специализация* предназначено для срабатывания в более частных ситуациях, чем соответствующее *правило-обобщение*. Любое *правило-специализация* имеет приоритет в срабатывании перед соответствующим *правилом-обобщением*.

Механизм интерпретации запроса на множестве правил указывает, какие правила и в каком порядке применять, а также какие описания считать истинными. Запрос представляет собой набор данных, который включает значения параметров пространства состояний объекта, спецификации располагаемых типов управляющих воздействий и ресурсов, необходимых для реализации воздействий [15].

3 Требования и задачи

Рассматриваемая сетевая М2М-система включает G производящих и S потребляющих узлов, соединенных L дугами.

В соответствии с физико-технологическими характеристиками изучаемой системы (при фиксированной схеме сети и составе производящего оборудования) потоки $y_i = \sum_g a_{ig}x_g + \sum_s b_{is}z_s$ ($i = 1, \dots, L$) по дугам линейно зависят от количества x_g производимого ($c_g \leq x_g \leq C_g$, $g = 1, \dots, G$, c_g и C_g — наименьшее и наибольшее допустимое значение для состава производящего оборудования, работающего в узле g) и прогнозируемого количества z_s потребляемого продукта ($d_s \leq z_s \leq D_s$, $s = 1, \dots, S$, d_s и D_s — наименьшее и наибольшее значение для количества продукта, потребляемого в узле s). Постановка и решение задачи поиска значений коэффициентов a_{ig} и b_{is} ($i = 1, \dots, L$, $g = 1, \dots, G$, $s = 1, \dots, S$) на основе результатов измерений были предложены в [16].

Набор требований, отражающих ограничения на пропускную способность каждой i -й дуги:

$$\left\{ \begin{array}{l} 0 \leq \sum_g a_{ig}x_g + \sum_s b_{is}z_s \leq Y_i [\leftarrow p_i]; \quad i = 1, \dots, L; \quad c_g \leq x_g \leq C_g; \\ g = 1, \dots, G; \quad d_s \leq z_s \leq D_s; \quad s = 1, \dots, S; \quad 0 < p_i \leq \infty \end{array} \right\}, \quad (1)$$

где Y_i — наибольшее допустимое значение потока продукта по i -й дуге; p_i — опциональный приоритет требования (при $p_i = \infty$ требование является обязательным, при иных значениях — *ориентирующим*).

Набор требований баланса в каждом g -м производящем узле, к которому примыкает m_g дуг с номерами $r = 1, \dots, m_g$:

$$\left\{ \begin{array}{l} -\delta_g \leq x_g - \sum_r \left(\sum_g a_{rg}x_g + \sum_s b_{rs}z_s \right) \leq \delta_g [\leftarrow p_g]; \quad c_g \leq x_g \leq C_g; \\ g = 1, \dots, G; \quad d_s \leq z_s \leq D_s; \quad s = 1, \dots, S; \quad 0 < p_g \leq \infty \end{array} \right\}, \quad (2)$$

где δ_g — допустимая (по технологическим основаниям) величина ошибки баланса в узле g ; p_g — опциональный приоритет требования для g -го производящего узла.

Набор требований баланса в каждом s -м потребляющем узле ($s = 1, \dots, S$), к которому примыкает m_s дуг с номерами $p = 1, \dots, m_s$:

$$\left\{ \begin{aligned} -\delta_s \leq \sum_p \left(\sum_g a_{pg} x_g + \sum_s b_{ps} z_s \right) - z_s \leq \delta_s [\leftarrow p_s]; \quad c_g \leq x_g \leq C_g; \\ g = 1, \dots, G; \quad d_s \leq z_s \leq D_s; \quad s = 1, \dots, S; \quad 0 < p_s \leq \infty \end{aligned} \right\}, \quad (3)$$

где δ_s — допустимая (по технологическим основаниям) величина ошибки баланса в узле s ; p_s — опциональный приоритет требования для s -го потребляющего узла.

◇ Себестоимость продукции q_g , производимой в g -м узле, контролируется цифровым двойником, который для каждого отрезка планирования на основе данных измерений строит зависимость себестоимости от нагрузки и представляет ее в виде линейной функции, используемой для формирования экономического показателя качества функционирования производящего узла: $q_g(x_g) = q_g^* + k_g x_g$ ($c_g \leq x_g \leq C_g$, $g = 1, \dots, G$; c_g и C_g — наименьшее и наибольшее допустимое значение для состава производящего оборудования, назначенного для работы в узле g на отрезке планирования; q_g^* — составляющая себестоимости, не зависящая от нагрузки узла). ◇

Требования (1)–(3) используются для формулировок задачи, зависящих от результатов анализа портретов ситуаций.

3.1 Уменьшение затрат на производство

Задача заключается в поиске вектора распределения производственных нагрузок $\mathbf{x} = (x_1, x_2, \dots, x_G)$ ($c_g \leq x_g \leq C_g$, $g = 1, \dots, G$, c_g и C_g — наименьшее и наибольшее допустимое значение для состава производящего оборудования, работающего в узле g), при котором достигаются наименьшие суммарные затраты на производство продукции $M(\mathbf{x}) = \sum_g q_g(x_g) [\leftarrow p_M]$ ($g = 1, \dots, G$), где p_M — опциональный приоритет требования уменьшения затрат на производство, при выполнении требований (1)–(3).

3.2 Уменьшение затрат на производство и доставку продукции

Найти вектор распределения нагрузок $\mathbf{x} = (x_1, x_2, \dots, x_G)$ ($c_g \leq x_g \leq C_g$, $g = 1, \dots, G$, c_g и C_g — наименьшее и наибольшее допустимое значение для состава производящего оборудования, работающего в узле g), при котором при выполнении (1)–(3) достигаются наименьшие общие затраты на производство и доставку продукции

$$Q(\mathbf{x}) = \sum_g q_g(x_g) + \sum_i q_i \sum_g a_{ig} x_g [\leftarrow p_Q] \quad (g = 1, \dots, G, \quad i = 1, \dots, L),$$

где p_Q — опциональный приоритет требования уменьшения общих затрат.

◇ Себестоимость q_i доставки единицы продукции по дуге i контролируется специальным цифровым двойником. ◇

4 Метод решения

Для решения задачи применен *метод целевого перемещения решения*, разработанный и программно реализованный А. В. Ильиным [17].

Для удобства читателя кратко воспроизведем здесь этот метод (дальнейшие обозначения не связаны с обозначениями разд. 3).

Начальная точка (первый вариант решения) может быть выбрана экспертом произвольно. По умолчанию предлагается компромиссное решение (чебышевская точка). Затем эксперт, проанализировав значения функций и оценив степень эффективности и реализуемости решения, может наложить на любые функции требования к их изменению. Величины изменений задают направление и дискрет перемещения в следующую точку. Кроме изменения можно задать и требование фиксации некоторой функции (когда значение функции устраивает эксперта и он хочет оставить его неизменным при дальнейшем перемещении). Таким образом эксперт пошагово приближается к решению, сочетающему, по его мнению, эффективность и реализуемость. Любое решение (т. е. всякая точка траектории движения) может быть занесено в базу возможных планов для последующего анализа.

Шаг целевого перемещения решения выполняется следующим способом. Пусть $\mathbf{x}' = (x'_1, \dots, x'_n)$ ($x'_j \geq 0, j = 1, \dots, n$) — решение, полученное на предыдущем шаге, и сформированы требования перемещения из \mathbf{x}' в целевую точку $\mathbf{x} = (x_1, \dots, x_n)$ ($x_j \geq 0, j = 1, \dots, n$): $\{F_i(\mathbf{x}) = F_i(\mathbf{x}') + h_i[p_i]\}$, где $F_i(\mathbf{x}) = a_{i1}x_1 + \dots + a_{in}x_n, i = 1, \dots, l; 0 < p_i \leq \infty$ для $h_i \neq 0, p_i = \infty$ для $h_i = 0$.

Формально система требований может быть несовместной. Поэтому требования с $h_i \neq 0$ рассматриваются как *ориентирующие*, а h_i — как *задающий шаг* (нередко отличающийся от реального шага, который может быть получен для данного набора требований). Если реальные и задающие шаги имеют один и тот же знак для всех требований, новая точка в любом случае повышает эффективность решения.

Точка \mathbf{x} отыскивается следующим образом. Сначала вычисляется проекция точки на гиперплоскость $a_{i1}x_1 + \dots + a_{in}x_n = F_i(\mathbf{x}') + h_i \forall h_i \neq 0$. Направляющий вектор нормали к этой гиперплоскости равен (a_{i1}, \dots, a_{in}) . Чтобы найти проекцию, необходимо дать переменным приращения $(ha_{i1}, \dots, ha_{in})$ (где h — неизвестное число).

Перемещение по нормали дает приращение функции $h(a_{i1}^2 + \dots + a_{in}^2)$, которое должно быть равно h , так что $h = h_i / (a_{i1}^2 + \dots + a_{in}^2)$ (естественно, $a_{i1}^2 + \dots + a_{in}^2 \neq 0$). Таким образом получается проекция $(x'_1 + a_{i1}h_1 / (a_{i1}^2 + \dots + a_{in}^2), \dots, (x'_n + a_{in}h_n / (a_{i1}^2 + \dots + a_{in}^2))$.

Когда проекции найдены $\forall h_i \neq 0$, ищутся требуемые приращения по всем переменным: $\Delta x_{ji} = a_{ij}h_i / (a_{i1}^2 + \dots + a_{in}^2)$ ($j = 1, \dots, n$; индексы функций пробегают значения от 1 до s , $1 \leq s \leq l$).

Далее вычисляется конец среднедействующей векторов-нормалей по формулам:

$$x_j = x'_j + \frac{p_1 \Delta x_{j1} + \dots + p_s \Delta x_{js}}{p_1 + \dots + p_s} \quad (j = 1, \dots, n).$$

Такая точка «ближе» к тем гиперплоскостям, которые соответствуют требованиям с более высокими приоритетами. Если приоритеты не были установлены, они полагаются равными 1 и получаются следующие формулы:

$$x_j = x'_j + \frac{\Delta x_{j1} + \dots + \Delta x_{js}}{s} \quad (j = 1, \dots, n).$$

Если $\exists k : h_k = 0$ ($1 \leq k \leq l$), конец среднедействующей проецируется на гиперплоскость $a_{k1}x_1 + \dots + a_{kn}x_n = F_k(x')$; в противном случае на роль искомой точки x претендует конец среднедействующей.

Затем контролируется неотрицательность переменных ($x_j \geq 0$, $j = 1, \dots, n$): отрицательные обнуляются.

И, наконец, проверяется совпадение знаков реальных и задающих шагов. Если они совпадают для всех требований, вычисленная точка считается целевой точкой x . Если не совпадают, эксперту предлагается скорректировать систему требований.

5 Заключение

1. Предложена методология ситуационного планирования производства (электроэнергии, воды и др.) в сетевой М2М-системе. Модель объекта описывает экономические характеристики производства и доставки продукции, а также *распределение потоков продукта по дугам сети* [16]. На каждом шаге планирования постановка задачи, составленная из *обязательных и ориентирующих требований*, зависит от результатов анализа *портретов ситуаций*. Основной составляющей вычислительного ядра служит *метод целевого перемещения решения*, предназначенный для решения линейных задач ситуационного планирования и управления [17].
2. Методология ориентирована на реализацию в виде *интернет-сервисов*, функционирующих на основе *цифровых двойников*, и рассчитана на применение в сетевых М2М-системах с динамично изменяющимися характеристиками объекта и экономического поведения производителей и потребителей.

Литература

1. Masud A. S., Hwang C. L. An aggregate production planning model and application of three multiple objective decision methods // Int. J. Prod. Res., 1980. Vol. 18. P. 741–752.

2. Foote B. L., Ravindran A., Lashine S. Production planning and scheduling: Computational feasibility of multi-criteria models of production, planning and scheduling // Comput. Ind. Eng., 1988. Vol. 15. P. 129–138.
3. Nam S. J., Logendran R. Aggregate production planning — a survey of models and methodologies // Eur. J. Oper. Res., 1992. Vol. 61. P. 255–272.
4. Byrne M. D., Bakir M. A. Production planning using a hybrid simulation — analytical approach // Int. J. Prod. Econ., 1999. Vol. 59. P. 305–311.
5. Hung Y. F., Shih C. C., Chen C. P. Evolutionary algorithms for production planning problems with setup decisions // J. Oper. Res. Soc., 1999. Vol. 50. P. 857–866.
6. Wang R. C., Fang H. H. Aggregate production planning with multiple objectives in a fuzzy environment // Eur. J. Oper. Res., 2001. Vol. 133. P. 521–536.
7. Leung S. C., Wu Y., Lai K. K. Multi-site aggregate production planning with multiple objectives: A goal programming approach // Prod. Plan. Control, 2003. Vol. 14. P. 425–436.
8. Ильин В. Д. Оперативное распределение нагрузок в энергосистеме // Автоматизированные системы контроля и управления. — Киев: Институт кибернетики АН УССР, 1973. С. 70–76.
9. Цифровая экономика Российской Федерации: Программа, утвержденная Распоряжением Правительства Российской Федерации от 28 июля 2017 г. № 1632-р. 87 с. <http://d-russia.ru/wp-content/uploads/2017/07/programma-tsifrov-econ.pdf>.
10. Machine to Machine (M2M) Communications Technical Report 802.16p-10/0005. IEEE, 2010.
11. Kim R. Y. Efficient wireless communications schemes for machine to machine communications // Comm. Com. Inf. Sc., 2011. Vol. 181. No. 3. P. 313–323.
12. Lien S. Y., Liao T. H., Kao C. Y., Chen K. C. Cooperative access class barring for machine-to-machine communications // IEEE T. Wirel. Commun., 2012. Vol. 11. No. 1. P. 27–32.
13. The Digital Twin. — General Electric, 2018. https://www.ge.com/digital/sites/default/files/The-Digital-Twin_Compressing-Time-to-Value-for-Digital-Industrial-Companies.pdf.
14. Ильин В. Д. Основания ситуационной информатизации. — М.: Наука, Физматлит, 1996. 180 с.
15. Ильин В. Д., Гавриленко Ю. В., Ильин А. В., Макаров Е. М. Математические средства ситуационной информатизации. — М.: Наука, Физматлит, 1996. 88 с.
16. Ильин А. В., Ильин В. Д. Модель распределения потоков в однопродуктовой сетевой системе // Системы и средства информатики, 2018. Т. 28. № 2. С. 170–177.
17. Ильин А. В. Экспертное планирование ресурсов. — М.: ИПИ РАН, 2013. 58 с.
18. Puyin A. V., Puyin V. D. Situational online resource planning in accordance with mandatory and orienting rules // Системы и средства информатики, 2018. Т. 28. № 1. С. 177–191.
19. Ильин В. Д. Символьное моделирование // Большая российская энциклопедия, 2018. http://dev.bigenc.ru/technology_and_technique/text/4010980.

Поступила в редакцию 14.09.18

SITUATIONAL PRODUCTION PLANNING IN A NETWORK M2M-SYSTEM

*A. V. Ilyin*¹ and *V. D. Ilyin*²

¹State Research Institute of Aviation Systems, 7 Viktorenko Str., Moscow 125319, Russian Federation

²Dorodnicyn Computing Center of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 40 Vavilova Str., Moscow 119333, Russian Federation

Abstract: A methodology for situational planning of production (electricity, water, etc.) in a network M2M (machine-to-machine) system is proposed. The methodology includes methods for constructing an object model using measurement data, forming portraits of situations (target, achieved, and starting), and interactive production planning. The object model describes economic characteristics of the nodes and distribution of the product streams moving along the network arcs, depending on the amount of product sent from the producing nodes and delivered to the consuming nodes. At each step of planning, the problem formulation depends on the results of analysis of the situations' portraits. The key component of the computational core is the method of target displacement of solution. The methodology is oriented to implementation by means of online services and is designed for use in network M2M-systems.

Keywords: network M2M-system; situational production planning; system of mandatory and orienting requirements; method of target displacement of solution; online service

DOI: 10.14357/08696527180413

References

1. Masud, A. S., and C. L. Hwang. 1980. An aggregate production planning model and application of three multiple objective decision methods. *Int. J. Prod. Res.* 18:741–752.
2. Foote, B. L., A. Ravindran, and S. Lashine. 1988. Production planning and scheduling: Computational feasibility of multi-criteria models of production, planning and scheduling. *Comput. Ind. Eng.* 15:129–138.
3. Nam, S. J., and R. Logendran. 1992. Aggregate production planning — a survey of models and methodologies. *Eur. J. Oper. Res.* 61:255–272.
4. Byrne, M. D, and M. A. Bakir. 1999. Production planning using a hybrid simulation — analytical approach. *Int. J. Prod. Econ.* 59:305–311.
5. Hung, Y. F, C. C. Shih, and C. P. Chen. 1999. Evolutionary algorithms for production planning problems with setup decisions. *J. Oper. Res. Soc.* 50:857–866.
6. Wang, R. C., and H. H. Fang. 2001. Aggregate production planning with multiple objectives in a fuzzy environment. *Eur. J. Oper. Res.* 133:521–536.
7. Leung, S. C., Y. Wu, and K. K. Lai. 2003. Multi-site aggregate production planning with multiple objectives: A goal programming approach. *Prod. Plan. Control* 14:425–436.

8. Ilyin, V. D. 1973. *Operativnoe raspredelenie nagruzok v energosisteme* [A methodology for operative load distribution in the power system]. *Avtomatizirovannyye sistemy kontrolya i upravleniya* [Automated control and monitoring systems]. Kiev: Institute of Cybernetics of the Academy of Sciences of Ukraine. 70–76.
9. Government of the Russian Federation. July 28, 2017. Tsifrovaya ekonomika Rossiyskoy Federatsii [The program “Digital Economy of the Russian Federation”]. Approved by Government Order No. 1632-r. 87 p. Available at: <http://d-russia.ru/wp-content/uploads/2017/07/programma-tsifrov-econ.pdf> (accessed September 10, 2018).
10. IEEE. 2010. Machine to Machine (M2M) Communications Technical Report 802.16p-10/0005.
11. Kim, R. Y. 2011. Efficient wireless communications schemes for machine to machine communications. *Comm. Com. Inf. Sc.* 181(3):313–323.
12. Lien, S. Y., T. H. Liao, C. Y. Kao, and K. C. Chen. 2012. Cooperative access class barring for machine-to-machine communications. *IEEE T. Wirel. Commun.* 11(1):27–32.
13. General Electric. 2018. The Digital Twin. Available at: https://www.ge.com/digital/sites/default/files/The-Digital-Twin_Compressing-Time-to-Value-for-Digital-Industrial-Companies.pdf (accessed October 31, 2018).
14. Ilyin, V. D. 1996. *Osnovaniya situatsionnoy informatizatsii* [Fundamentals of situational informatization]. Moscow: Nauka, Fizmatlit. 180 p.
15. Ilyin, V. D., Yu. V. Gavrilenko, A. V. Ilyin, and E. M. Makarov. 1996. *Matematicheskie sredstva situatsionnoy informatizatsii* [Mathematical means of situational informatization]. Moscow: Nauka, Fizmatlit. 88 p.
16. Ilyin, A. V., and V. D. Ilyin. 2018. Model’ raspredeleniya potokov v odnoproductovoy setevoy sisteme [The model of flows distribution in a single-product network system]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(2):170–177.
17. Ilyin, A. V. 2013. *Ekspertnoe planirovanie resursov* [Expert resource planning]. Moscow: IPI RAN. 58 p. Available at: <https://www.researchgate.net/publication/259654243> (accessed September 10, 2018).
18. Ilyin, A. V., and V. D. Ilyin. 2018. Situational online resource planning in accordance with mandatory and orienting rules. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(1):177–191.
19. Ilyin, V. D. 2018. Simvol’noe modelirovanie [Symbolic modeling]. *Bol’shaya rossiyskaya entsiklopediya* [Great Russian Encyclopedia]. Available at: http://dev.bigenc.ru/technology_and_technique/text/4010980 (accessed September 10, 2018).

Received September 14, 2018

Contributors

Ilyin Alexander V. (b. 1975) — Candidate of science (PhD) in technology, leading engineer, State Research Institute of Aviation Systems, 7 Viktorenko Str., Moscow 125319, Russian Federation; ilyin@res-plan.com

Ilyin Vladimir D. (b. 1937) — Doctor of Science in technology, professor, leading scientist, Dorodnicyn Computing Center of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 40 Vavilova Str., Moscow 119333, Russian Federation; vdilyin@yandex.ru

ЭЛЕМЕНТЫ САМООБУЧЕНИЯ В СИСТЕМЕ ИЗВЛЕЧЕНИЯ БИОГРАФИЧЕСКИХ ФАКТОВ Т-ПАРСЕР

И. М. Адамович¹, О. И. Волков²

Аннотация: Статья посвящена дальнейшему развитию системы автоматического извлечения фактов из текстов историко-биографической направленности Т-парсер — составной части технологии автоматизации историко-биографического исследования. Намечаются пути увеличения скорости обработки за счет использования механизма самообучения. Описываются и обосновываются применяемые формы самообучения, формулируются возможные проблемы. Приводится классификация ветвлений при разборе текстов. Описывается и обосновывается механизм фильтрации при формировании базы данных (БД) прецедентов на основе методов статистического приемочного контроля по альтернативному признаку. Приводится описание обновленного алгоритма парсинга и экспериментальной проверки его эффективности по сравнению с предыдущей версией, осуществленной на реальных текстах историко-биографической направленности. Приводятся результаты экспериментов, подтверждающие высокую эффективность обновленного алгоритма и его применимость в технологии автоматизации историко-биографического исследования, предназначенной для широкого круга не являющихся профессиональными историками и биографами пользователей, что актуально в связи со все увеличивающимся общественным интересом к семейной истории.

Ключевые слова: извлечение фактов из текстов; самообучение; ветвления; статистический приемочный контроль; обучающая выборка

DOI: 10.14357/08696527180414

1 Введение

В статье [1] описана технология автоматизации историко-биографического исследования. Важной составной частью данной технологии является система автоматического извлечения фактов из текстов историко-биографической направленности Т-парсер [2]. Несмотря на то что на данный момент существуют весьма развитые программные средства, предназначенные для решения задач автоматической обработки текстов и семантического анализа их смыслового содержания [3, 4], специфика историко-биографического исследования [5] потребовала создания оригинального средства извлечения биографических фактов из текстов с поддержкой ряда возможностей, позволяющих использовать ее в технологии поддержки историко-биографических исследований:

¹Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, Adam@amsd.com

²Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, Volkov@amsd.com

- разрешение анафорических связей, когда интерпретация некоторого выражения зависит от другого выражения, ранее встречавшегося в тексте;
- поддержка древовидной структуры биографического факта [6];
- ориентированность на многопользовательский распределенный режим работы в соответствии с технологией автоматизации историко-биографического исследования [1].

T-парсер был реализован на базе основных принципов GLR (generalized left-to-right) алгоритма [7] и продемонстрировал высокую эффективность выделения фактов: F-мера [8] на тестовых примерах составила от 80% до 84% [2] при приемлемой скорости обработки текста. Проведенная модификация алгоритма [9], основанная на исключении циклов из грамматики и ее линейного упорядочения, позволила значительно сократить время обработки, за счет чего появилась возможность применить механизм разветвления цепочек разбора и их параллельную обработку при возникновении неоднозначности с целью отбрасывания ошибочных цепочек на последующих этапах. Эта модификация дополнительно повысила качество распознавания, выраженное в значениях F-меры, но привела к некоторому увеличению времени обработки. В результате возникла потребность в проведении дальнейшей модификации алгоритма за счет внедрения в него элементов самообучения, что позволяет значительно скомпенсировать увеличение времени обработки с сохранением высокого качества распознавания.

2 Разветвление цепочек разбора

Механизм разветвления цепочек разбора для их параллельной обработки при возникновении неоднозначности реализован следующим образом: в соответствии с линейным порядком на множестве правил [9] на каждом шаге обработки к фразе применяются только правила с соответствующим уровнем и статусом. Успешные применения правил к фразе описываются множеством $M = \{(n_i, s_i)\}$ пар, состоящих из номера правила и его сдвига — номера позиции во фразе, начиная с которой применилось правило [2]. Для элементов M будем называть их областями применения позиции фразы, начиная со сдвига, к которым было применено правило. Элементы M будем называть пересекающимися, если пересекаются их области применения. На фразе выявляются области пересечения S_i — идущие подряд позиции фразы, в которых для каждого элемента M , область применения которого содержится в S_i , существует иной элемент M , пересекающийся с ним. Если S — область пересечения, то множество его правил $M(S)$ определим как подмножество M , все элементы которого имеют области применения, принадлежащие S . Вариантом $V(S)$ для области пересечения S будем называть такое подмножество $M(S)$, все элементы которого не пересекаются. Если вариант $V(S) = \{(n_i, s_i)\}$, то его характеристикой назовем упорядоченный по сдвигу перечень его правил, т. е. $\text{Наг}(V) = \{n_i\}$. Размером покрытия $\text{Size}(V)$ варианта V будем называть сумму длин областей применения всех его элементов. Размер

покрытия задает отношение частичного порядка μ на множестве вариантов для данной области пересечения: $\mu(V) = \text{Size}(V)$.

Для каждой области пересечения S формируется ее набор ветвлений $\text{Br}(S) = \{V_i(S) \mid \exists j : \mu(V_j(S)) > \mu(V_i(S))\}$, каждому элементу которого соответствует ветвь разбора. Описателем ветвления $V(S) \in \text{Br}(S)$ будем называть совокупность $(\{\text{Наг}(V_i(S) \mid V_i(S) \in \text{Br}(S))\}, n \mid V_n = V(S), \text{Status})$, где Status — код причины признания данной ветви успешной или ошибочной в результате дальнейшего разбора по следующему алгоритму: на каждом шаге работы парсера для каждой ветви сравниваются размеры покрытия на всей фразе для всех ветвей одной области. Если будет выявлена ветвь с меньшим размером покрытия по сравнению с другой ветвью, то данная ветвь признается ошибочной и ее дальнейший разбор останавливается. Статус соответствующего ветвления принимается «неуспешным». Если после полного завершения разбора несколько ветвей завершились с одинаковым размером покрытия, то в качестве результирующей из них выбирается произвольная, а остальные назначаются ошибочными и их статус принимается «условно неуспешным». Статус результирующей ветви принимается «успешным».

3 Самообучение

Поскольку, несмотря на высокую скорость разбора, обеспеченную введением линейного порядка на множестве правил [9], параллельная обработка нескольких ветвей является достаточно времязатратным способом определения оптимальной ветви, целесообразно в алгоритм ввести механизм самообучения, с тем чтобы некоторые типовые ситуации ветвления определялись в результате обработки некоторой обучающей выборки текстов и далее уже отбрасывались на этапе формирования набора вариантов для области пересечения.

Для организации самообучения были выбраны следующие принципы [10].

1. Обучение осуществляется без учителя, что означает, что для каждого прецедента требуемое решение заранее не известно и система должна сформировать его самостоятельно. В данном случае это достигается разбором всех ветвей и нахождением среди них успешной.
2. В результате обучения формируется система, основанная на прецедентах, БД которой содержит описания не обобщенных ситуаций, а собственно сами ситуации или прецеденты.
3. Но при этом система формируется без присущего системам, основанным на прецедентах, механизма нечеткого поиска с получением множества допустимых альтернатив, каждая из которых оценивается некоторым коэффициентом уверенности. Механизм формирования решения об отбрасывании варианта базируется на принципах построения индуктивных систем, т. е. анализ новой ситуации сводится к выбору прецедента, который полностью определяет эту ситуацию.

Следует учитывать, что алгоритмам, построенным на указанных принципах, присущи следующие проблемы [11].

1. Большой объем обучающей выборки приводит к значительному росту времени обучения, а малый объем дает менее надежные результаты и увеличивает вероятность так называемого переобучения, когда алгоритм хорошо обрабатывает примеры из обучающей выборки, но относительно плохо работает на примерах, не участвовавших в обучении.
2. Проблема противоречивости данных, когда объекты с одним и тем же признаковым описанием могут иметь разные исходы, не решаемая корректно без применения вероятностных методов обучения.
3. Проблема индуктивных систем, состоящая в том, что индукция может давать из истинных посылок ложное заключение.

Таким образом, корректность выбранных принципов самообучения применительно к задачам, решаемым системой Т-парсер, требует экспериментальной проверки.

4 Модификация алгоритма

Для реализации вышеуказанных принципов целесообразно провести следующую модификацию алгоритма.

1. В процессе обучения при определении ошибочных ветвей необходимо записывать их описатель ветвления в БД прецедентов. Данная процедура должна выполняться не для всех ветвлений, а только для их некоторого подкласса, разбор ветвлений из которого всегда завершается однозначным образом. Для определения этого подкласса должна быть проведена классификация ветвлений.
2. Необходимо следующим образом модифицировать отношение порядка μ на множестве вариантов: если $\text{Size}(V_1) = \text{Size}(V_2)$, но ветвление, соответствующее варианту V_1 , присутствует в БД прецедентов и помечено как ошибочное, а для V_2 это не так, то $\mu(V_1) < \mu(V_2)$.

5 Классификация видов ветвлений

Анализ ветвлений при обработке текстов историко-биографической направленности показывает, что ветвления вызываются следующими причинами.

1. Неоднозначность русского языка: ветвления, вызванные этой причиной, будут интерпретироваться Т-парсером по-разному в разных ситуациях и, соответственно, порождать как успешные, так и неуспешные ветви.
2. Недостаточная точность процедуры линейного упорядочения грамматики [9], в результате которой правилу с фактически более высоким приоритетом назначается тот же уровень, что и правилу с фактически более низким приоритетом. Ветвления, вызванные этой причиной, будут всегда одинаково

интерпретироваться Т-парсером в любых ситуациях и, соответственно, породить только успешные или только неуспешные ветви.

3. Избыточность грамматики, что неизбежно при одновременном независимом создании правил выделения фактов множеством не связанных между собой исследователей. Ветвления, вызванные этой причиной, будут всегда одинаково интерпретироваться Т-парсером, и порожденные этой причиной ветви никогда не будут отбрасываться как ошибочные. Статус таких ветвей будет всегда «успешным» или «условно неуспешным».

Первая категория должна исключаться из обучающей выборки, поскольку, как указано выше, не может корректно обрабатываться без применения вероятностных методов. Вторая и третья категории должны быть включены в обучающую выборку, поскольку обрабатываются однозначно. Поэтому для формирования БД прецедентов должны быть выработаны критерии, с высокой вероятностью отсекающие ветвления первой категории.

6 Принципы фильтрации ветвлений при формировании базы данных прецедентов

Для успешного формирования БД прецедентов необходима процедура надежной фильтрации ветвлений, позволяющая отсеять ветвления, вызванные неоднозначностью языка и приводящие для одного и того же описателя ветвления в разных ситуациях к различным интерпретациям, т. е. к различным вариантам выбора результирующей ветви. Отсечение должно проводиться по ограниченной выборке наблюдений, сформированной в рамках обработки обучающей выборки, не позволяющей в силу своего объема выявить все факты различной интерпретации каждого описателя ветвления. Данная задача сходна с задачей статистического приемочного контроля по альтернативному признаку. Статистический приемочный контроль — это контроль качества конечной генеральной совокупности изделий (партии), проводимый на основе выборок ограниченного объема. Он основан на применении методов математической статистики [12]. Применительно к БД прецедентов применение методов статистического контроля означает:

- под изделием понимаем ветвление с данным описателем;
- под партией изделий объемом N понимаем совокупность всех ветвлений с данным описателем, возникших в процессе историко-биографического исследования;
- в совокупности объемом $n + 1$ всех ветвлений с данным описателем, возникших в процессе разбора обучающей выборки, первое ветвление будем называть определяющим, а остальные — выборкой объемом n ;
- под дефектом ветвления будем понимать результат его интерпретации, противоречащий результату интерпретации определяющего ветвления;
- приемочное число A_c (партия принимается, если число дефектных единиц продукции в выборке меньше или равно A_c , и бракуется, если больше A_c)

- принимая равным нулю, поскольку наличие даже одного дефекта в выборке для данного описателя ветвления однозначно относит его к первому типу;
- под уровнем дефектности партии будем понимать долю неверных интерпретаций для ветвлений с данным описателем за весь период проведения историко-биографического исследования; данный показатель должен соответствовать показателям качества распознавания Т-парсера в целом, а поскольку F-мера для Т-парсера превышает 80%, браковочный уровень годности q_m целесообразно установить равным $1 - F \approx 20\%$;
 - значения для риска поставщика α (вероятности того, что партия может быть ошибочно забракована) и риска потребителя β (вероятности того, что партия может быть ошибочно принята) принимаем обычными для задач подобного рода: $\alpha = 0,05$ и $\beta = 0,10$ [13].

Поскольку объем выборки мал и доля дефектных изделий также мала, для определения объема выборки при одноступенчатом контроле может быть использовано распределение Пуассона [12]:

$$\sum_{i=0}^{A_c} \frac{pn^i e^{-pn}}{i!} = \beta. \quad (1)$$

Для $A_c = 0$ формула (1) после логарифмирования принимает вид:

$$n = -\frac{\ln \beta}{q_m}.$$

Поскольку $\beta = 0,1$ и $q_m = 0,2$, минимальный объем выборки n_{\min} составляет 12 элементов.

Применительно к проблеме формирования БД прецедентов это означает, что ветвления, описатели которых были зафиксированы при обработке обучающей выборки менее $n_{\min} + 1$ раз (с учетом определяющего ветвления), должны отсекаются, поскольку по ним нельзя сделать достоверного вывода о принадлежности их к первой категории ветвлений.

7 Алгоритм формирования базы данных прецедентов

Поскольку грамматика может состоять из нескольких независимых наборов правил, для каждого такого набора проводится следующая процедура.

1. По результатам обработки обучающей выборки текстов формируется набор описателей всех произошедших ветвлений.
2. Поскольку условно неуспешный статус назначается произвольно, в наборе описателей могут присутствовать описатели ветвлений с одинаковой характеристикой варианта, но с противоположными статусами для ветвей: неуспешной ветви первого ветвления соответствует успешная ветвь второго, а успешная ветвь первого — условно неуспешной второго. В этом случае

корректируются описатели второго ветвления: ветвям, соответствующим неуспешным ветвям первого ветвления, устанавливается неуспешный статус, а ветви, соответствующей успешной ветви первого ветвления, устанавливается успешный статус.

3. Для каждого уникального описателя подсчитывается число его вхождений в набор.
4. Описатели с числом вхождений не более n_{\min} отбрасываются, а прочие помещаются в БД прецедентов.

8 Экспериментальная проверка модифицированного алгоритма

Как следует из приведенного выше списка проблем, возможно вызванных использованием механизмов самообучения, проверке подлежат следующие факты:

- отсутствие «проблемы переобучения», т. е. сравнимость результатов обработки обучающей и тестовой выборки;
- отсутствие проблемы противоречивости данных, т. е. корректность критерия формирования БД прецедентов;
- отсутствие ошибок обработки, возникающих за счет опоры на принципы построения индуктивных систем, т. е. принципиальная применимость описанного алгоритма самообучения для задач автоматического извлечения фактов из текстов историко-биографического содержания на естественном языке.

Первая задача решается за счет сравнения результатов обработки обучающей и тестовой выборки по такому показателю, как относительное снижение времени обработки за счет использования БД прецедентов.

Вторая и третья задачи решаются за счет сравнения результатов обработки текстов с использованием и без использования БД прецедентов по показателям качества распознавания.

В качестве обучающей выборки был взят набор фрагментов текстов историко-биографического содержания общим объемом около 430 КБ. Обработка выявила в данном наборе 402 случая ветвления, которым соответствовали 69 уникальных описателей ветвления. Из этих описателей только у 8 число ветвлений превысило n_{\min} , но совокупное число ветвлений для них — 268, что составляет 66,6% от общего числа ветвлений. Парсинг данного набора в режиме тестирования показал снижение времени обработки на 12,4% по сравнению с парсингом в режиме обучения.

Поскольку расчет качества распознавания, выраженного в значении F-меры, требует ручного выделения фактов из обрабатываемого текста для последующего сравнения результатов с результатами парсинга, а на больших объемах текста эта процедура является весьма трудоемкой, было принято решение опереться на известное значение показателей качества для немодифицированного алгоритма и ограничиться сравнением перечней автоматически извлеченных фактов для

обоих режимов. В результате обработки обучающей выборки без использования БД прецедентов (в режиме обучения) и обработки с использованием БД прецедентов (в режиме тестирования) были сформированы идентичные структуры из 2460 деревьев фактов с 3739 листьями (атомарными фактами) [6]. Таким образом, различий в качестве распознавания при обработке обучающей выборки в режимах с использованием и без использования БД прецедентов не выявлено.

В качестве тестовой выборки был взят набор текстов историко-биографического содержания общим объемом около 1770 КБ. Обработка выявила в данном наборе 1321 случай ветвления, которым соответствовали 77 уникальных описателей ветвления. Из этих описателей только у 19 число ветвлений превысило n_{\min} , но совокупное число ветвлений для них — 981, что составляет 74,3% от общего числа ветвлений. Парсинг данного набора в режиме тестирования показал снижение времени обработки на 10,1% по сравнению с парсингом в режиме обучения.

По вышеописанным причинам сравнение качества распознавания на тестовой выборке в режимах с использованием и без использования БД прецедентов также производилось сравнением перечней автоматически извлеченных фактов для обоих режимов. В результате обработки тестовой выборки без использования БД прецедентов (в режиме обучения) была сформирована структура из 8440 деревьев фактов с 12 826 листьями (атомарными фактами). В результате ее обработки с использованием БД прецедентов (в режиме тестирования) была сформирована структура из 8439 деревьев фактов с 12 821 листьями (атомарными фактами). Структуры отличаются на 1 дерево фактов с 5 листьями, что составляет менее 0,04%. Таким образом, различия в качестве распознавания при обработке тестовой выборки в режимах с использованием и без использования БД прецедентов можно считать несущественными.

В результате можно констатировать практически полное совпадение результатов обработки текстов с использованием и без использования БД прецедентов по показателям качества распознавания и близость (с учетом степени насыщенности текста точками ветвления) результатов обработки обучающей и тестовой выборок по такому показателю, как относительное снижение времени обработки за счет использования БД прецедентов.

9 Выводы

Проверка показала, что проведенная модификация алгоритма системы Т-парсер за счет внедрения в него элементов самообучения позволяет значительно снизить время обработки с сохранением высокого качества автоматического извлечения фактов из текстов историко-биографического содержания на естественном языке.

Система Т-парсер может быть беспрепятственно использована в технологии автоматизации историко-биографического исследования, предусматривающей

одновременное независимое создание правил выделения фактов множеством не связанных между собой исследователей.

Литература

1. *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.
2. *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.
3. *Артемов М. А., Владимиров А. Н., Селезнев К. Е.* Обзор систем анализа естественного текста на русском языке // Вестник Воронежского государственного университета. Сер. Системный анализ и информационные технологии, 2013. № 2. С. 189–194.
4. *Будзко В. И., Калинин Ю. П., Козеренко Е. Б., Хорошилов Ал-др А., Хорошилов Ал-ей А.* Машинная грамматика русского языка // Системы высокой доступности, 2017. Т. 13. № 3. С. 19–33.
5. *Адамович И. М., Волков О. И.* Средства поддержки интернет-поиска при проведении биографических исследований // Системы и средства информатики, 2014. Т. 24. № 2. С. 178–192. doi: 10.14357/08696527140212.
6. *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.
7. *Погорелов Д. А., Таразанов А. М., Волкова Л. Л.* От LR к GLR: обзор синтаксических анализаторов // Новые информационные технологии в автоматизированных системах, 2017. № 20. С. 245–250.
8. *Агеев М. С., Кураленок И. Е., Некрестьянов И. С.* Официальные метрики РОМИП'2010 // Российский семинар по оценке методов информационного поиска: Труды РОМИП'2010. — Казань, 2010. С. 172–187.
9. *Адамович И. М., Волков О. И.* Линейное упорядочение множества правил в системе извлечения биографических фактов Т-парсер // Системы и средства информатики, 2018. Т. 28. № 3. С. 217–226. doi: 10.14357/08696527180317.
10. *Тельнов Ю. Ф.* Интеллектуальные информационные системы. — М.: МЭСИ, Московская высшая банковская школа, 2001. 118 с.
11. *Ветров Д. П., Кропотов Д. А.* Байесовские методы машинного обучения. — М.: МГУ им. М. В. Ломоносова, 2007. 133 с.
12. *Быков Ю. М.* Статистический приемочный контроль по альтернативному признаку. — Волгоград: ВолгГТУ, 2002. 52 с.
13. ГОСТ Р 50779.75-99 Статистические методы. Последовательные планы выборочного контроля по альтернативному признаку. — М.: Изд-во стандартов, 2000. 45 с.

Поступила в редакцию 15.05.18

ELEMENTS OF MACHINE LEARNING IN THE T-PARSER SYSTEM OF FACTS EXTRACTION

I. M. Adamovich and O. I. Volkov

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

Abstract: The article focuses on the further development of the system of facts automatic extraction from historical texts T-parser which is a component of the technology of historical and biographical research automation. The article outlines the ways to increase the parsing speed by using machine learning. The chosen forms of machine learning are described and reasoned and the possible problems are formulated. The classification of parsing bifurcations is given. The mechanism of filtering for the precedent database creation based on the methods of statistical quality control on an alternative basis is described and reasoned. The description of the updated parsing algorithm and experimental verification of its effectiveness in comparison with the previous version carried out with real historical texts are adduced. The results of experiments which confirm high efficiency of the updated algorithm and its applicability to the technology of historical and biographical research automation are described. The technology is intended for a broad range of nonprofessional users, which is topical with regard to the increasing public interest to family history.

Keywords: facts extraction from texts; machine learning; bifurcation; statistical quality control; training set

DOI: 10.14357/08696527180414

References

1. Adamovich, I. M., and O. I. Volkov. 2016. Tekhnologiya raspredelennogo avtomatizirovannogo analiza istoricheskikh tekstov [The distributed automated technology of historical texts analysis]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 26(3):148–161. doi: 10.14357/08696527160311.
2. 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.
3. Artemov, M. A., A. N. Vladimirov, and K. E. Seleznyov. 2013. Obzor sistem analiza estestvennogo teksta na russkom yazyke [Review on natural language analysis systems for Russian language]. *Proceedings of Voronezh State University. Ser. Systems Analysis and Information Technologies* 2:189–194.
4. Budzko V. I., Yu. P. Kalinin, E. B. Kozerenko, A. A. Khoroshilov, and A. A. Khoroshilov. 2017. Mashinnaya grammatika russkogo yazyka [Machine grammar

- of the Russian language]. *Sistemy vysokoy dostupnosti* [Highly Available Systems] 13(3):19–33.
5. Adamovich, I. M., and O. I. Volkov. 2014. Sredstva podderzhki internet-poiska pri provedenii biograficheskikh issledovaniy [The technology of Internet search as a part of biographic investigation]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 24(2):178–192. doi: 10.14357/08696527140212.
 6. Adamovich, I. M., and O. I. Volkov. 2016. Ierarkhicheskaya forma predstavleniya biograficheskogo fakta [Hierarchical format of biographical fact]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 26(2):108–122. doi: 10.14357/08696527160207.
 7. Pogorelov, D. A., A. M. Tarzanov, and L. L. Volkova. 2017. Ot LR k GLR: obzor sintaksicheskikh analizatorov [From LR to GLR: Overview of parsers]. *Novye informatsionnye tekhnologii v avtomatizirovannykh sistemakh* [New Information Technologies Automated Systems] 20:245–250.
 8. Ageev, M. S., I. E. Kuralenok, and I. S. Nekrest'yanov. 2010. Ofitsial'nye metriki ROMIP 2010 [ROMIP'2010 official metrics]. *Rossiyskiy seminar po otsenke metodov informatsionnogo poiska: Trudy ROMIP'2010* [ROMIP: Russian Information Retrieval Evaluation Seminar Proceedings]. Kazan: Kazan University Publ. 172–187.
 9. Adamovich, I. M., and O. I. Volkov. 2018. Lineynoe uporyadochivanie mnozhestva pravil v sisteme izvlecheniya biograficheskikh faktov T-parser [Linear ordering of rules set in the system of facts extraction T-parser]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(3):217–226. doi: 10.14357/08696527180317.
 10. Tel'nov, U. F. 2001. *Intellektual'nye informatsionnye sistemy* [Intelligent information systems]. Moscow: Moscow State University of Economics, Statistics, and Informatics, “Moscow Bank Higher School” Institute 118 p.
 11. Vetrov, D. P., and D. A. Kropotov. 2007. *Bayesovskie metody mashinnogo obucheniya* [Bayesian methods in machine learning]. Moscow: MGU Publ. 133 p.
 12. Bykov, Yu. M. 2002. *Statisticheskii priemochnyy kontrol' po al'ternativnomu priznaku* [Statistical quality control on an alternative basis]. Volgograd: VSTU. 52 p.
 13. GOST R 50779.75-99. 2000. *Statisticheskie metody. Posledovatel'nye plany vyborochnogo kontrolya po al'ternativnomu priznaku* [Statistical methods. Sequential sampling plans for inspection by attributes]. Moscow: Standardinform Publ. 45 p.

Received May 15, 2018

Contributors

Adamovich Igor M. (b. 1934)— Candidate of Science (PhD) in technology, leading scientist, Institute of Informatics Problem, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; Adam@amsd.com

Volkov Oleg I. (b. 1964) — leading programmer, Institute of Informatics Problem, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; Volkov@amsd.com

НАДКОРПУСНАЯ БАЗА ДАННЫХ КОННЕКТОРОВ: РАЗВИТИЕ СИСТЕМЫ ТЕРМИНОВ ПРОЕКТИРОВАНИЯ*

И. М. Зацман¹, М. Г. Кружков²

Аннотация: Рассматривается задача развития системы терминов, ориентированного на проектирование надкорпусной базы данных (НБД) коннекторов, которая представляет собой новую категорию информационных лингвистических ресурсов. Надкорпусная база данных содержит параллельные выровненные тексты, в которых одному или нескольким предложениям оригинальных текстов поставлены в соответствие одно или несколько предложений их переводов. Надкорпусные базы данных предназначены для аннотирования широкого спектра языковых единиц (ЯЕ), но в данной статье рассматривается задача аннотирования коннекторов. Характерная черта НБД состоит в том, что их ориентация на решение задач аннотирования привела к появлению новых сущностей и понятий в компьютерной лингвистике. Для их описания в начале 2017 г. была предложена система терминов. С одной стороны, эти термины используются лингвистами для представления новых фундаментальных знаний, полученных ими в процессе аннотирования и контрастивного исследования языковых единиц. С другой стороны, они применяются при разработке архитектурных решений НБД, их информационного, лингвистического и программного обеспечения. За время, прошедшее с момента первого описания системы терминов, существенно расширился спектр задач, решаемых с помощью НБД. Этим обусловлена необходимость развития этой системы терминов, ориентированного на проектирование НБД, чему и посвящена данная статья.

Ключевые слова: надкорпусная база данных; система терминов; аннотирование языковых единиц; параллельные тексты; корпусная лингвистика; коннекторы

DOI: 10.14357/08696527180415

1 Введение

Необходимость исследования ЯЕ с использованием больших объемов параллельных выровненных текстов стала основным мотивом создания НБД [1–6]. Для описания задач и функций НБД в 2017 г. была предложена система терминов [7]. С одной стороны, эта система используется лингвистами для представления новых фундаментальных знаний, полученных ими в процессе аннотирования

*Работа выполнена при поддержке РФФИ (проект 16-24-41002).

¹Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, izatsman@yandex.ru

²Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, magnit75@yandex.ru

и контрастивного исследования ЯЕ. С другой стороны, она применяется при разработке архитектурных решений НБД, их информационного, лингвистического и программного обеспечения.

За время, прошедшее с момента публикации первой версии системы терминов в начале 2017 г., существенно расширился спектр задач, решаемых с помощью НБД. Этим обусловлена необходимость развития этой системы, ориентированного на проектирование НБД. Описание предлагаемого ее развития и является основной целью статьи. В ней новая (вторая) версия системы терминов иллюстрируется на примере описания структурных схем НБД коннекторов.

Надкорпусная база данных представляет собой реляционную базу данных на основе SQL Server с веб-интерфейсом, благодаря чему она доступна пользователям через большинство веб-браузеров и может поддерживать работу всех ее пользователей одновременно. Чтобы описать структурные схемы НБД коннекторов, сначала опишем новую систему терминов, применяемых в процессе ее проектирования, используя в качестве основы первую версию системы [7]. Затем будут описаны сами структурные схемы НБД коннекторов и связи между ее полями.

2 Термины, лежащие в основе проектирования надкорпусной базы данных

Отметим, что система терминов НБД эволюционировала в процессе становления новой категории лингвистических ресурсов, получивших название НБД (термин введен в работе [4]), и продолжает эволюционировать. Поэтому приведенные здесь термины не всегда совпадают с терминами, использованными ранее в других работах [2, 5–7]. Описывая термины, будем по возможности указывать, какие варианты названия каждого понятия использовались ранее.

Основная идея концепции НБД заключается в том, что *языковые единицы* некоторого языка исследуются с помощью НБД «в зеркале» одного или нескольких других языков на материале параллельных текстов [8]. Язык, ЯЕ которого выступают основным предметом исследования, будем называть *базовым языком*, а языки, в зеркале которых исследуются ЯЕ базового языка, будем называть *референтными языками*. Обычно с помощью НБД исследуется определенный класс ЯЕ базового языка, который будем называть *исходным классом ЯЕ* (соответственно, ЯЕ, входящие в этот класс, будем называть *исходными ЯЕ*). В частности, в рамках проекта РФФИ № 16-24-41002 с помощью НБД исследуются коннекторы русского языка. Таким образом, в этом проекте базовым языком является русский язык, а исходным классом ЯЕ служат коннекторы русского языка. Референтным языком в рамках данного проекта выступает французский язык.

Контрастивное исследование ЯЕ базового языка в зеркале некоторого референтного языка (или языков) на материале параллельных текстов может осуществляться двумя способами анализа.

Прямой анализ. При рассмотрении переводов с базового языка на некоторый референтный язык исследователей интересует, какие ЯЕ референтного языка в переводе соответствуют исходным ЯЕ базового языка в оригинале. При такой схеме анализа базовый язык является *языком оригинала*, а референтный язык выступает *языком перевода*. Выявленные лингвистами соответствия между исходными ЯЕ базового языка и ЯЕ референтного языка, которые называются *прямыми моделями перевода*, представляют собой кортеж, т. е. упорядоченную пару вида ⟨ЯЕ базового языка; ЯЕ референтного языка⟩. В прямой модели перевода ЯЕ референтного языка является *результатом перевода*. Таким образом, при прямом анализе исходные ЯЕ базового языка исследуются в зеркале результатов их перевода на некоторый референтный язык. В табл. 1 язык оригинала — русский, язык перевода — французский, ЯЕ базового языка — *если... так ведь*, ЯЕ референтного языка — *si*, модель перевода — ⟨если... так ведь; si⟩.

Таблица 1 Русско-французский перевод

Оригинальный текст	Перевод
— <i>Боже мой!</i> — говорил Обломов. — Да если слушать Штольца, так ведь до тетки век дело не дойдет! [И. А. Гончаров. Обломов (1848–1859)]	— <i>Mon Dieu!</i> dit Oblomov. <i>Mais si</i> on <i>écoutait Stolz, on n'en arriverait jamais</i> <i>à la tante!</i> [Перевод. Luba Jurgenson (1988)]

Реверсивный анализ. При рассмотрении переводов с некоторого референтного языка на базовый язык исследователей интересует, какие ЯЕ референтного языка в оригинале вызывают появление исходных ЯЕ базового языка в переводе. При таком анализе референтный язык является *языком оригинала*, а базовый язык — *языком перевода*. Выявленные при таком анализе соответствия между ЯЕ референтного языка и ЯЕ базового языка, которые называются *реверсивными моделями перевода*, представляют собой кортеж вида ⟨ЯЕ референтного языка; ЯЕ базового языка⟩, а ЯЕ референтного языка служат *стимулом перевода*. Таким образом, при реверсивном анализе исходные ЯЕ базового языка исследуются в зеркале их стимулов перевода с некоторого референтного языка. В табл. 2 язык оригинала — французский, язык перевода — русский, ЯЕ референтного языка — *et*, ЯЕ базового языка — *а... ведь*, модель перевода — ⟨et; а... ведь⟩.

Следует отметить также не всегда очевидное различие между терминами *ЯЕ* и *употребление ЯЕ*, которое является принципиальным для проектирования НБД. С одной стороны, можно говорить о некоторой ЯЕ безотносительно к какому-либо контексту ее употребления в конкретном тексте. С другой стороны, можно говорить об *употреблениях ЯЕ* в определенных контекстах в рамках конкретных текстов.

Это важное различие до сих пор недостаточно четко проводилось в отдельных статьях, посвященных концепции НБД, из-за чего эти понятия могли смешиваться. Это различие является ключевым в процессе проектирования НБД. Если

Таблица 2 Французско-русский перевод

Оригинальный текст	Перевод
<p><i>A qui s'étonnait qu'un tel vol eût pu s'accomplir aussi facilement, le sous-gouverneur Gauthier Ralph se bornait à répondre qu'à ce moment même, le caissier s'occupait d'enregistrer une recette de trois shillings six pence, et qu'on ne saurait avoir l'œil à tout.</i> [Жюль Верн. Le tour du monde en quatre-vingt jours (1872)]</p>	<p><i>В ответ на удивленные вопросы, как могла произойти подобная кража, помощник управляющего банком Готье Ральф ограничивался следующим ответом: «В эту минуту кассир вписывал в приход поступление в три шиллинга и шесть пенсов, а за всем ведь не уследишь».</i> [Перевод. Н. Габинский (1939)]</p>

в НБД надо хранить некоторую ЯЕ и модели ее перевода, то описание этой ЯЕ будет включать, как правило, несколько десятков записей. Если же надо хранить употребления некоторой ЯЕ и результаты их перевода, то описания этих употреблений будут включать сотни или даже тысячи записей.

Рассмотрим, например, коннектор *то есть*, который в НБД является одной из исследуемых ЯЕ. Ниже приводится несколько случаев употребления этой ЯЕ в четырех разных произведениях:

- (1) С Разумихиным же он почему-то сошелся, *то есть* не то что сошелся, а был с ним сообщительнее, откровеннее. [Ф. М. Достоевский. Преступление и наказание (1866)]
- (2) Он худошав; щек у него почти вовсе нет, *то есть* есть кость да мускул. . . [И. А. Гончаров. Обломов (1859)]
- (3) Доктора не могли определить. *То есть* определяли, но различно. [Л. Н. Толстой. Смерть Ивана Ильича (1886)]
- (4) У меня в доме живет и теща, *то есть* мать моей жены, и дети. . . [Н. В. Гоголь. Нос (1836)]

Таким образом, различие между понятиями ЯЕ и *употребление ЯЕ* аналогично различию между понятиями *слово* и *словоупотребление*. Далее в этой главе, когда будет важно провести различие между этими двумя понятиями, собственно ЯЕ будем кратко обозначать !ЯЕ!, а *употребление ЯЕ* — #ЯЕ#.

Отметим также, что в отдельных статьях по тематике НБД вместо широкого термина ЯЕ, используемого в этой главе, используются более узкие по значению термины. Например, в большинстве ранних статей вместо этого использовалось понятие лексико-грамматическая форма (ЛГФ), а в НБД коннекторов часто используется термин *речевая реализация* (РР) коннектора.

Основным информационным объектом НБД выступает *переводное соответствие* (ПС) (в ряде предшествующих работ вместо термина ПС использовался

термин *моноэквиваленция* (МЭ)). Переводные соответствия в общем случае представляют собой упорядоченные пары текстовых фрагментов (кортежи) вида $\langle \text{FO}; \text{FT} \rangle$, включающих употребления исследуемых ЯЕ и их переводы. Вторым фрагментом (FT) является переводом первого фрагмента (FO). Предполагается, что в каждом кортеже эти фрагменты семантически и функционально соответствуют друг другу (хотя это соответствие может быть и неполным).

Как уже отмечалось, НБД проекта РФФИ № 16-24-41002 обеспечивает лингвистическое аннотирование и исследование коннекторов русского языка. В связи с этим в рамках этой НБД одна из двух составляющих кортежа $\langle \text{FO}; \text{FT} \rangle$ (либо фрагмент оригинала, либо фрагмент перевода) включает в себя i -е употребление некоторого русского коннектора ЯЕ, т. е. $\# \text{ЯЕ}(i) \#$, $i = 1, \dots, N$, где N — число употреблений этого коннектора, аннотированных в НБД. В прямой модели перевода $\# \text{ЯЕ}(i) \#$ включена во фрагмент оригинала (пример 5), а в реверсивной модели — во фрагмент перевода (пример 6).

Другая составляющая кортежа $\langle \text{FO}; \text{FT} \rangle$ (это также может быть либо фрагмент оригинала, либо фрагмент перевода) является фрагментом на референтном (французском) языке. Обычно он включает в себя r -е употребление некоторой ЯЕ референтного языка, т. е. $\# \text{ЯЕ}(r) \#$, $r = 1, \dots, M$, где M — число употреблений этой ЯЕ, аннотированных в НБД¹. В данном контексте $\# \text{ЯЕ}(r) \#$ выступает соответствием, или эквивалентом, употребления исходной $\# \text{ЯЕ}(i) \#$.

В прямой модели перевода $\# \text{ЯЕ}(r) \#$ соответствует фрагменту перевода (см. пример 5), а при использовании реверсивной модели $\# \text{ЯЕ}(r) \#$ соответствует фрагменту оригинала (см. пример 6).

- (5) Теоретически, буфет обслуживал сотрудников Дома прессы. *В том числе* — журналистов из многотиражек. [Сергей Довлатов. Чемодан (1986)]

En théorie, ce buffet avait pour clients les membres de la maison de la Presse. *Donc*, les journalistes. [Перевод. Jacques Michaut-Paterno (2001)]

Здесь фрагмент оригинала включает i -е употребление ЯЕ *в том числе* (принадлежит исходному классу русских коннекторов), а фрагмент перевода — r -е употребление французской ЯЕ *donc*, которая в данном контексте является результатом перевода i -го употребления ЯЕ *в том числе* на французский язык.

- (6) Cette lutte l'avait fatigué *comme s'il avait combattu*, à l'exemple de Jacob, contre un ange. [Honoré de Balzac. L'Élixir de longue vie (1830)]

¹Строго говоря, необходимо было бы указывать два индекса для обозначения каждого употребления, а именно: $\# \text{ЯЕ}(j, i) \#$, где j — номер исследуемой ЯЕ, i — номер ее употребления, и $\# \text{ЯЕ}(p, r) \#$, где p — номер эквивалента исследуемой ЯЕ, r — номер его употребления. Здесь при описании употреблений ЯЕ и их эквивалентов первые индексы будем опускать, чтобы упростить обозначения. Отметим, что в этих обозначениях N зависит от индекса j , так как у разных ЯЕ разное число употреблений в общем случае, т. е. $i = 1, \dots, N(j)$, а M — от p , так как у разных эквивалентов также может быть разное число употреблений, т. е. $r = 1, \dots, M(p)$.

Этот поединок его утомил, *как будто* он, наподобие Иакова, боролся с ангелом. [Перевод. Б. Грифцов (1930–1950)]

Здесь фрагмент перевода включает i -е употребление *ЯЕ как будто* (принадлежит исходному классу русских коннекторов), а фрагмент оригинала — r -е употребление французской *ЯЕ comme si*, которая в данном контексте выступает стимулом i -го употребления *ЯЕ как будто* в переводе на русский язык.

Таким образом, учитывая различие между понятиями *ЯЕ* и *употребление ЯЕ*, в прямом направлении перевода ПС следует представлять в виде $\langle \text{FO}(\# \text{ЯЕ}(i)\#); \text{FT}(\# \text{ЯЕ}(r)\#) \rangle$, где $\text{FO}(\# \text{ЯЕ}(i)\#)$ — фрагмент оригинала, включающий i -е употребление некоторой *ЯЕ*, а $\text{FT}(\# \text{ЯЕ}(r)\#)$ — фрагмент перевода, включающий r -е употребление французской *ЯЕ* (см. пример 5), тогда как в реверсивном направлении перевода ПС должна иметь вид $\langle \text{FO}(\# \text{ЯЕ}(r)\#); \text{FO}(\# \text{ЯЕ}(i)\#) \rangle$ (см. пример 6).

Фрагменты оригинала и перевода, входящие в ПС, аннотируются с помощью наборов признаков-тегов. Во-первых, для каждого фрагмента проставляется *главный признак*, который указывает, что данный фрагмент следует рассматривать как случай употребления некоторой *ЯЕ* соответствующего языка. Для базового языка набор главных признаков — это множество *ЯЕ* базового языка, зарегистрированных в НБД в процессе аннотирования: $\{! \text{ЯЕ}(j)!, j = 1, \dots, K\}$, где K — общее число таких *ЯЕ*. В НБД проекта РФФИ № 16-24-41002 — это коннекторы русского языка.

Для референтных языков набор главных признаков — это множество зарегистрированных в НБД *ЯЕ* референтного языка, которые являются либо результатами, либо стимулами перевода для элементов множества $! \text{ЯЕ}(j)!$ базового языка: $\{! \text{ЯЕ}(p)!, p = 1, \dots, L\}$, где L — общее число таких *ЯЕ*¹. В процессе пополнения НБД наборы главных признаков по мере необходимости расширяются, особенно когда речь идет о наборах главных признаков референтных языков. Таким образом, K и L могут увеличиваться, так как НБД коннекторов поддерживает возможность расширения множества $\{! \text{ЯЕ}(j)!\}$, если лингвисты решат добавить в него новые *ЯЕ*.

Подытоживая сказанное, повторим, что *главный признак*, присваиваемый фрагменту, входящему в ПС, указывает на то, какая именно *ЯЕ* в этом фрагменте выступает объектом рассмотрения. Это может быть: (1) $! \text{ЯЕ}(j)!$ исходного класса базового языка, (2) $! \text{ЯЕ}(p)!$ референтного языка или (3) один из признаков-тегов, который указывает на то, что главный признак выделить не удастся [9].

Помимо главных признаков в процессе лингвистического аннотирования используются *дополнительные признаки*. Всю совокупность дополнительных признаков, характеризующих разные аспекты употребления *ЯЕ*, можно представить как *фасетную классификацию* [6]. При этом семантически связанные наборы

¹По состоянию на 30.08.2018 для прямого направления перевода (с русского на французский) $K = 1029$ и $L = 1456$.

значений дополнительных признаков (в НБД они называются *кластерами дополнительных признаков*) можно рассматривать как фасеты, а сами значения дополнительных признаков — как рубрики этих фасетов.

Достоинство представления рубрик в форме фасетной классификации состоит в том, что путем комбинации рубрик различных фасетов можно осуществлять поиск коннекторов заданного типа, обладающих некоторым набором значений признаков и встречающихся в определенном контексте [10]. При этом в работе с ПС имеется возможность комбинировать главные и дополнительные признаки для двух языков для некоторого направления перевода, что еще больше расширяет поисковый функционал НБД.

3 Структурные схемы надкорпусной базы данных

В данном разделе описана структура НБД коннекторов, которая обеспечивает решение лингвистических задач проекта. Ее можно «разбить» на две части: корпусную и надкорпусную. Корпусная часть предназначена для хранения параллельных текстов, полученных из Национального корпуса русского языка (НКРЯ) [11]. На материале текстов НКРЯ осуществляется исследование исходного множества $\{ЯЕ(j), j = 1, \dots, K\}$. В этой части БД содержится информация о произведениях, составляющих корпуса, переводах этих произведений, выровненных парах фрагментов оригинала и перевода. Это дает возможность представить исполнителям проекта параллельные тексты в выровненном виде (см. табл. 1 и 2). Также в корпусной части содержится информация о словах, входящих во фрагменты оригинала и перевода, и о результатах морфологического анализа этих слов. На основе этой информации в НБД осуществляется первичный поиск фрагментов, в которых могут находиться интересующие исследователей ЯЕ. Логическая структурная схема корпусной части НБД представлена на рис. 1¹.

В надкорпусной части НБД содержится информация о ПС и входящих в них употреблениях исследуемых ЯЕ (примеры 5 и 6). Логическая структурная схема надкорпусной части НБД приведена на рис. 2.

4 Заключение

Как видно из рис. 2, структурная схема надкорпусной части НБД состоит из нескольких взаимосвязанных таблиц. Входящие в ПС фрагменты текста базового языка и соответствующие им фрагменты текстов референтных языков описываются с помощью одной и той же таблицы «Фрагмент». Кроме того, таблица «Фрагмент» используется для описания употреблений ЯЕ оригинала и перевода. Это позволяет упростить и унифицировать работу с коннекторами в прямых и реверсивных переводных соответствиях.

¹На рис. 1 приводятся обобщенные диаграммы таблиц, полей и связей НБД, поскольку их детальные варианты тяжелы для восприятия.

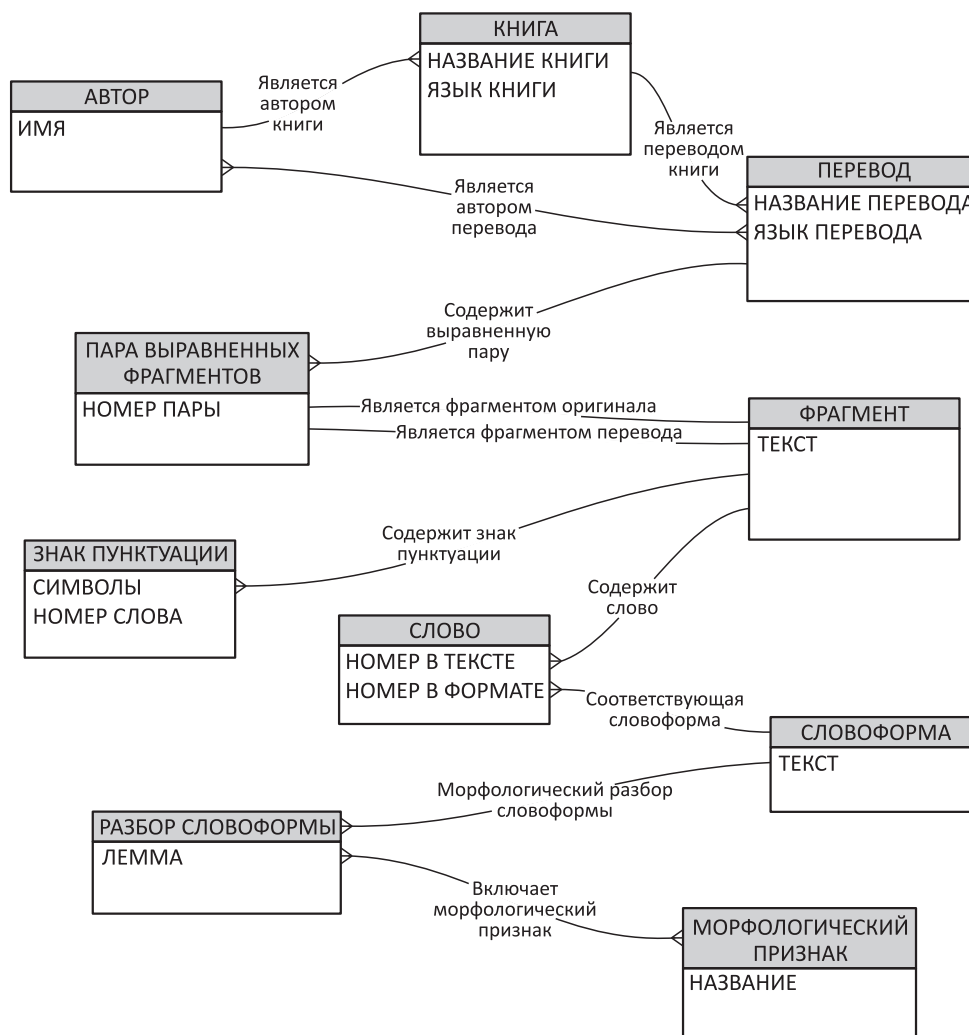


Рис. 1 Логическая структурная схема корпусной части НБД. Прямоугольники соответствуют таблицам НБД, линии между прямоугольниками — связям между таблицами. Связи один-к-одному обозначаются простыми линиями, связи один-ко-многим — линиями с одним простым и одним разветвленным концом, связи многие-ко-многим — линиями с двумя разветвленными концами (например, одной словоформе может соответствовать несколько морфологических разборов, которые могут относиться к разным леммам)

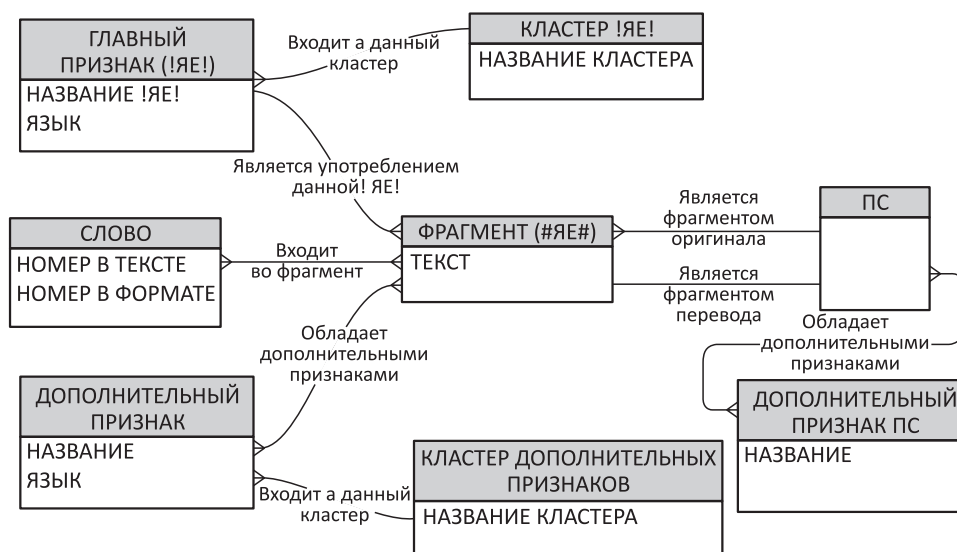


Рис. 2 Логическая структурная схема надкорпусной части НБД

Связь с корпусной частью НБД осуществляется через таблицу «Слово»: каждый текстовый фрагмент, входящий в ПС, определяется как совокупность употреблений коннекторов в текстах некоторого произведения или его перевода, что позволяет задавать необходимые исследователю границы для контекстов употреблений $\#ЯЕ(j, i)\#$ и $\#ЯЕ(p, r)\#$. Это нужно при описании ЯЕ, для которых требуется задавать широкий контекст, выходящий за рамки выровненной пары предложений в параллельном тексте, что важно при описании коннекторов, которые нередко могут связывать несколько соседних предложений.

Литература

1. Добровольский Д. О., Кретов А. А., Шаров С. А. Корпус параллельных текстов: архитектура и возможности использования // Национальный корпус русского языка: 2003–2005. — М.: Индрик, 2005. С. 263–296.
2. Loiseau S., Sitchinava D. V., Zalizniak Anna A., Zatsman I. M. Information technologies for creating the database of equivalent verbal forms in the Russian–French multivariant parallel corpus // Информатика и её применения, 2013. Т. 7. Вып. 2. С. 100–109.
3. Сичинава Д. В. Использование параллельного корпуса для количественного изучения лингвоспецифичной лексики // Язык, литература, культура: актуальные проблемы изучения и преподавания. — М.: МАКС ПРЕСС, 2014. Вып. 10. С. 37–44.

4. *Кружков М. Г.* Информационные ресурсы контрастивных лингвистических исследований: электронные корпуса текстов // Системы и средства информатики, 2015. Т. 25. № 2. С. 140–159.
5. *Зализняк Анна А., Зацман И. М., Инькова О. Ю., Кружков М. Г.* Надкорпусные базы данных как лингвистический ресурс // Корпусная лингвистика-2015: Тр. 7-й Междунар. конф. — СПб.: СПбГУ, 2015. С. 211–218.
6. *Зацман И. М., Инькова О. Ю., Кружков М. Г., Попкова Н. А.* Представление кроссязыковых знаний о коннекторах в надкорпусных базах данных // Информатика и её применения, 2016. Т. 10. Вып. 1. С. 106–118.
7. *Зализняк Анна А., Зацман И. М., Инькова О. Ю.* Надкорпусная база данных коннекторов: построение системы терминов // Информатика и её применения, 2017. Т. 11. Вып. 1. С. 100–106.
8. *Зализняк Анна А.* База данных межъязыковых эквиваленций как инструмент лингвистического анализа // Компьютерная лингвистика и интеллектуальные технологии: По мат-лам ежегодной междунар. конф. «Диалог». — М.: Изд-во РГГУ, 2016. Вып. 15(22). С. 854–866.
9. *Зацман И. М., Кружков М. Г., Лощилова Е. Ю.* Методы анализа частотности моделей перевода коннекторов и обратимость генерализации статистических данных // Системы и средства информатики, 2017. Т. 27. № 4. С. 164–176.
10. *Inkova O. Yu., Popkova N. A.* Statistical data as information source for linguistic analysis of Russian connectors // Информатика и её применения, 2017. Т. 11. Вып. 3. С. 123–131.
11. Национальный корпус русского языка. <http://www.ruscorpora.ru>.

Поступила в редакцию 07.09.18

SUPRACORPORA DATABASE OF CONNECTIVES: DESIGN-ORIENTED EVOLUTION OF THE TERM SYSTEM

I. M. Zatsman and M. G. Kruzhkov

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

Abstract: This article examines the process of design-oriented evolution of the term system for supracorpora databases (SCDB) which represent a new category of information resources in linguistics. The SCDB is based on parallel texts, i. e., texts placed alongside their translations and aligned with them at the sentence level. Although SCDBs are designed for annotation of a wide variety of linguistic items and their correspondences, this article specifically considers annotation of connectives. The annotation-centered design of SCDBs has led to emergence of new entities and notions in computer linguistics, and in the beginning of 2017, a custom term system was proposed for them. On one hand, the proposed terms are used by linguists in order to describe new knowledge generated as a result of annotation and investigation of linguistic units. On

the other hand, these terms serve as a basis for design of the SCDB architecture and the associated dataware, lingware, and software. Since the first description of the terminology, the range of tasks accomplished with SCDBs has expanded significantly; hence, there is the need to further develop the initial design-oriented term system.

Keywords: supracorpora databases; term systems; annotation of linguistic units; parallel texts; corpus linguistics; connectives

DOI: 10.14357/08696527180415

Acknowledgments

This research was performed in the Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, with financial support of the Russian Foundation for Basic Research (grant No. 16-24-41002).

References

1. Dobrovol'skiy, D. O., A. A. Kretov, and S. A. Sharov. 2005. Korpus parallel'nykh tekstov: arkhitektura i vozmozhnosti ispol'zovaniya [Corpus of parallel texts: Architecture and applications]. *Natsional'nyy korpus russkogo yazyka: 2003–2005* [The Russian National Corpus 2003–2005]. Moscow: Indrik. 263–296.
2. Loiseau, S., D. V. Sitchinava, Anna A. Zalizniak, and I. M. Zatsman. 2013. Information technologies for creating the database of equivalent verbal forms in the Russian–French multivariant parallel corpus. *Informatika i ee Primeneniya — Inform. Appl.* 7(2):100–109.
3. Sichinava, D. V. 2014. Ispol'zovanie parallel'nogo korpusa dlya kolichestvennogo izucheniya lingvospetsifichnoy leksiki [Using a parallel corpus for quantitative research of language-specific lexis]. *Yazyk, literatura, kul'tura: aktual'nye problemy izucheniya i prepodavaniya* [Language, literature, culture: Urgent problems of research and teaching]. Moscow: MAKS Press. 10:37–44.
4. Kruzchkov, M. G. 2015. Informatsionnye resursy kontrastivnykh lingvisticheskikh issledovaniy: elektronnye korpusa tekstov [Information resources for contrastive studies: Electronic text corpora]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 25(2):140–159.
5. Zaliznyak, Anna A., I. M. Zatsman, O. Yu. In'kova, and M. G. Kruzchkov. 2015. Nadkorpurnye bazy dannykh kak lingvisticheskiy resurs [Supracorpora databases as a linguistic resource]. *7th Conference (International) on Corpus Linguistics Proceedings*. St. Petersburg: St. Petersburg State University. 211–218.
6. Zatsman, I. M., O. Yu. In'kova, M. G. Kruzchkov, and N. A. Popkova. 2016. Predstavlenie krossyazykovykh znaniy o konnektorakh v nadkorpurnykh bazakh dannykh [Representation of cross-lingual knowledge about connectors in supracorpora databases]. *Informatika i ee Primeneniya — Inform. Appl.* 10(1):106–118.

7. Zaliznyak, Anna A., I. M. Zatsman, and O. Yu. In'kova. 2017. Nadkorpurnaya baza dannykh konnektorov: postroyeniye sistemy terminov [Supracorpora database of connectives: Term system development]. *Informatika i ee Primeneniya — Inform. Appl.* 11(1):100–106.
8. Zaliznyak, Anna A. 2016. Baza dannykh mezh'yazykovykh ekvivalentsey kak instrument lingvisticheskogo analiza [A database of cross-linguistic equivalences as an instrument of linguistic analysis]. *Computational Linguistics and Intellectual Technologies: Conference (International) "Dialog" Proceedings*. Moscow: RGGU. 15(22):854–866.
9. Zatsman, I. M., M. G. Kruzhkov, and E. Yu. Loshchilova. 2017. Metody analiza chastotnosti modeley perevoda konnektorov i obratimost' generalizatsii statisticheskikh dannykh [Reversibility and alternativeness of generalization of connectives translation models in parallel texts]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 27(4):164–176.
10. Inkova, O. Yu., and N. A. Popkova. 2017. Statistical data as information source for linguistic analysis of Russian connectors. *Informatika i ee Primeneniya — Inform. Appl.* 11(3):123–131.
11. *Natsional'nyy korpus russkogo yazyka* [The Russian National Corpus]. Available at: <http://www.ruscorpora.ru> (accessed September 4, 2018).

Received September 7, 2018

Contributors

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

Kruzhkov Mikhail G. (b. 1975) — 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; magnit75@yandex.ru

МЕТОД ОПИСАНИЯ СТРУКТУРЫ НЕОДНОСЛОВНЫХ КОННЕКТОРОВ В НАДКОРПУСНЫХ БАЗАХ ДАННЫХ*

О. Ю. Инькова¹, М. Г. Кружков²

Аннотация: Представлен новый информационно-компьютерный метод описания структуры однословных коннекторов, реализованный в рамках Надкорпусной базы данных (НБД) коннекторов. На сегодняшний день структура коннекторов мало изучена, отсутствуют критерии для определения линейных границ коннекторов и их компонентов. В основе предлагаемого метода лежит когнитивно-семантический подход, при котором однословные коннекторы считаются более или менее свободными сочетаниями. При аннотировании коннекторов предлагается использовать двухуровневую фасетную классификацию, где аннотированию подлежат, с одной стороны, употребления коннекторов в тексте (контекстное аннотирование), с другой стороны — собственно внутренняя структура коннекторов (структурное аннотирование). Структурное аннотирование осуществляется по двум основаниям: определяются структурный тип и структурные составляющие коннекторов. Предлагаемая схема аннотации позволяет реализовать систему «перекрестных» кластеров, значительно расширяющую поисковые и статистические возможности НБД коннекторов. Кроме того, данный метод позволяет избегать субъективности при аннотировании однословных коннекторов в электронных лингвистических корпусах и заполнять лакуны в лингвистических знаниях: например, получать новые данные о сочетаемостных возможностях коннекторов русского языка.

Ключевые слова: надкорпусные базы данных; компьютерная обработка естественного языка; коннекторы; структура языковых единиц; вариативность языковых единиц; корпусная лингвистика; аннотирование; фасетная классификация

DOI: 10.14357/08696527180416

1 Введение

В связи с развитием корпусной лингвистики и систем обработки естественного языка перед специалистами в области лингвистики и информатики встает задача, связанная с распознаванием и интерпретацией логических связей между ситуациями, описанными в тексте. Согласно определению, данному в [1], языковые единицы, маркирующие логико-семантические отношения (ЛСО) между фраг-

*Работа выполнена при финансовой поддержке РФФИ (проект 16-06-00070).

¹Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, olyainkova@yandex.ru

²Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, magnit75@yandex.ru

ментами текста, описывающими ситуации, называются коннекторами. Многие коннекторы являются сочинительными или подчинительными союзами, однако их морфологическая природа далеко не ограничивается одной этой частью речи — функцию коннекторов могут выполнять единицы различных морфологических и синтаксических классов, среди которых можно выделить наречия и частицы (*иначе, также, кстати*), некоторые предлоги (*после*), предикативные структуры (*Дело в том, что...; Стоит... как...*).

Значительное число коннекторов состоит более чем из одного слова (*а также, но не только, по крайней мере* и пр.) — такие коннекторы будем называть *неоднословными*. Проблема описания структуры неоднословных коннекторов до последнего времени не являлась предметом специального обсуждения, структура их мало изучена, отсутствуют критерии для определения линейных границ неоднословных коннекторов и их компонентов. Более того, отсутствуют четкие критерии для определения того, какие речевые реализации следует считать едиными языковыми единицами, а какие — составными языковыми единицами. Например, в академической грамматике русского языка ([2], далее — РГ-80) речевая реализация *а значит* считается сочетанием союза с конкретизатором, в то время как *а также* — единой языковой единицей. Кроме того, многие коннекторы характеризуются высокой степенью формальной вариативности, что отражается и в их представлении в словарях и грамматиках. Например, в предметном указателе РГ-80 фиксируются такие союзы: *не то что... но* (*а; просто; даже; даже не*), *не только... но и* (*а и; но даже; а еще, а к тому же*). Заметим, что указанные варианты далеко не исчерпывают все сочетаемостные возможности формантов *не то что* и *не только* (подробнее см. [3–5]).

Цель этой работы состоит в том, чтобы представить новый информационно-компьютерный метод описания структуры неоднословных коннекторов, реализованный в рамках НБД коннекторов. Надкорпусная база данных (термин введен в [6]) представляет собой реляционную базу данных на основе системы управления базами данных SQL Server и позволяет формировать переводные соответствия для различных языковых единиц, в том числе неоднословных коннекторов, а также аннотировать эти единицы с использованием классификационных схем, которые пользователи-лингвисты могут настраивать с учетом своих потребностей. В разд. 2 будут представлены существующие подходы к аннотированию неоднословных коннекторов как для русского, так и для других языков. В разд. 3 будет рассмотрена общая структура НБД. В разд. 4 будет показано, как структурная схема НБД коннекторов была спроектирована применительно к задаче описания структуры неоднословных коннекторов в рамках когнитивно-семантического подхода.

2 Существующие подходы к описанию неоднословных коннекторов

Разработчики существующих аннотированных корпусов русского языка, в которых ставится вопрос о неоднословных языковых единицах (Хельсинский ан-

нотированный корпус (ХАНКО) [7] и Национальный корпус русского языка (НКРЯ) <http://ruscorpora.ru/obgrams.html>), исходят из того, что неоднослов-ные коннекторы являются устойчивыми сочетаниями, и на этом основании они задаются фиксированными списками, размеры которых в значительной степени разнятся. Так, для аннотирования неоднословных служебных слов создателями корпуса ХАНКО был разработан — на основе толковых словарей и словарей служебных слов русского языка — список служебных неоднословных единиц. Рубрика «Союзы» включает 323 единицы. Разработчики НКРЯ создали «Корпусный словарь неоднословных лексических единиц», которые квалифицируются как «устойчивые лексические обороты». Необходимость создания такого словаря объясняется тем, что морфологическая разметка дается в НКРЯ для орфографического слова, т. е. отделяемого пробелом. В основу словаря положены данные частотных коллокаций НКРЯ с дополнениями из словарей. Список оборотов в функции союза и союзного слова включает 159 единиц (т. е. меньше половины списка ХАНКО), для каждой из которых дано ориентировочное число употреблений. Заметим, что наиболее полный на сегодняшний день список союзов и союзных слов РГ-80 включает 666 единиц.

Кроме того, критерий частотности применяется в «Корпусном словаре» НКРЯ непоследовательно. В число «частотных» попадают обороты *добро бы еще* и *диви б* с единичными вхождениями в НКРЯ; фиксируются как самостоятельные лексические единицы *добро бы* и *добро б*, *если бы* и *если б*, тогда как из комбинаций с *да* фиксируются только *да и* и *да еще*, хотя поиск в НКРЯ по *да притом еще* выдает 57 вхождений, по *да вдобавок еще* — 50. Наконец, в список НКРЯ, в отличие от списка ХАНКО, включаются только коннекторы, составляющие которых расположены контактно, т. е. не разделены другими словами текста. Таким образом, вне списка остаются двухкомпонентные коннекторы (*если... то...*, *как... так и...*, *не только... но и...*), а также коннекторы с дистантным расположением элементов (*да ему и не нужно; но я вообще особый человек*).

Такой подход приводит к определенной бессистемности и субъективности в описании состава служебных единиц русского языка, поскольку критерий для внесения в список той или иной языковой единицы не определен с достаточной четкостью. В конечном итоге для разметки и составления «полных» списков служебных неоднословных языковых единиц используются словники существующих словарей.

В этой связи следует упомянуть пионерскую для своего времени работу [8]. В ней все эти вопросы ставятся с предельной четкостью, но не всегда получают ответ, в частности, потому, что это исследование было написано в «докорпусную» эпоху, а его авторы работали с картотекой примеров, включающей не более тысячи употреблений различных коннекторов. В сводной таблице, содержащей результаты исследования в том, что касается неоднословных коннекторов, представлены сочетаемостные возможности коннекторов русского языка, и одним из параметров описания является как раз возможность дистантного расположе-

ния составляющих неоднословных коннекторов. В таблицу, по словам авторов, включены «только самые активные» в образовании неоднословных коннекторов составляющие, а в самой таблице в силу ограниченности материала незаполненным оказывается большинство клеток.

«Списочный» подход используется также для других языков, где создаются корпуса, посвященные анализу дискурсивных отношений (ср., например, The Penn Discourse Treebank, <https://www.seas.upenn.edu/~pdtb>, French Discourse Treebank [9], ANNODIS, <http://redac.univ-tlse2.fr/corpus/annodis>). Хорошо известны списки показателей дискурсивных отношений, претендующие на исчерпывающие и создающиеся для автоматического распознавания и разметки текстов: список для английского языка первоначально содержал 200 показателей [10], а затем был доведен до 400 [11]; список для испанского языка включает 577 базовых форм, а с учетом их морфологических вариантов — 784 формы (см. его описание и критерии отбора в [12]); список для французского языка (LEXCONN, <http://www.linguist.univ-paris-diderot.fr/~croze/D/Lexconn.xml>), составленный как первый этап создания French Discourse Treebank, содержит 328 показателей ЛСО, хотя кандидатов изначально было около 600 [13]. При этом вопрос о формальной вариативности коннекторов не обсуждается, хотя списки показателей и составляются по гнездовому принципу: например, в списке LEXCONN в один «кластер» попадают формы *quand*, *quand même*, *quand bien même*, *quand bien même que*¹ и т. д.

Однако внимание в работах этого направления, ставящего своей задачей автоматический анализ текста, а не его моделирование, уделяется преимущественно самим отношениям, а списки коннекторов составляются для того, чтобы с их помощью машина могла идентифицировать то или иное дискурсивное отношение.

3 Структура надкорпусной базы данных коннекторов

Для описания структуры неоднословных коннекторов предлагается использовать когнитивно-семантический подход, при котором они считаются более или менее свободными сочетаниями, порождаемыми говорящим не только благодаря его знанию языка, но и в соответствии с его коммуникативными намерениями. Этим и объясняется тот факт, что некоторые, в значительной степени очевидные, ЛСО могут не выражаться никаким коннектором в тексте, а другие, значимые с коммуникативной точки зрения, могут выражаться довольно «тяжеловесными» формами (ср. *едва* по сравнению с *едва только... как тут же*). Основной единицей анализа при когнитивно-семантическом подходе становится *речевая реализация* (РР), т. е. та форма, в которой коннектор встречается в данном кон-

¹ Буквальный перевод, отражающий структуры этих языковых единиц: *quand* ‘когда’, *quand même* ‘когда даже’, *quand bien même* ‘когда же даже’, *quand bien même que* ‘когда же даже что’. Последние три единицы являются показателями уступительных отношений и эквивалентны русским коннекторам *однако* и *даже если*.

Таблица 1 Фрагмент выровненного русско-французского параллельного текста (Михаил Булгаков. Белая Гвардия, 1926, перевод на французский — Claude Ligny, 1970)

Исходный текст	Текст перевода
— Ну что ты, Алеша. . . лежи, молчи. . .	— Voyons, Aliocha. . . Allons, reste couché et tais-toi. . .
Пальто-то этой дамы у нас пока будет?	Le manteau de cette dame, on le garde ici, pour l'instant?
— Да, да. Чтобы Николка не вздумал тащить его.	— Oui, oui. Et que Nikolka ne s'avise pas de le porter.
А то на улице. . . Слышишь?	Sinon, dans la rue. . . Tu comprends?
Вообще, ради бога, не пускай его никуда.	En général, pour l'amour du ciel, ne le laisse aller nulle part.

клетном высказывании [14]. Именно она регистрируется в НБД, которая была спроектирована для того, чтобы дать возможность лингвистам применить для изучения структуры коннекторов и ее вариативности когнитивно-семантический подход.

Исходным материалом для НБД служат русско-французские параллельные тексты НКРЯ, которые включают в себя морфологическую разметку и прочие релевантные признаки. Каждый параллельный текст включает в себя один исходный текст и один или несколько текстов перевода, которые выровнены друг с другом на уровне предложений (см. пример в табл. 1).

Назначение НБД состоит в том, чтобы устанавливать переводные соответствия между текстами оригинала и перевода на уровне конкретных языковых единиц, а именно коннекторов, в том числе неоднословных. Кроме того, НБД обеспечивает возможность аннотирования единиц, входящих в переводные соответствия, с помощью специализированных наборов признаков (тегов). В конечном счете структура НБД позволяет осуществлять поиск и статистический анализ переводных соответствий между коннекторами русского языка и соответствующими им языковыми единицами французского языка¹.

На рис. 1 показана основная часть концептуальной схемы той части НБД коннекторов, которая отвечает за описание переводных соответствий. Из данной схемы видно, что в терминах НБД переводное соответствие представляет собой объединение фрагмента текста оригинала, заключающего в себе аннотируемый коннектор, и фрагмент текста перевода. Каждый из этих фрагментов имеет привязку к набору слов из корпуса, причем эти слова имеют различную маркировку или «статус», которые определяют, являются ли они частью данной языковой единицы или же функционально значимы для данной языковой единицы в дан-

¹В рамках данной статьи для простоты изложения будем рассматривать лишь русско-французское направление перевода, хотя НБД также охватывает русско-итальянские тексты и переводы на русский язык с французского и итальянского языков.

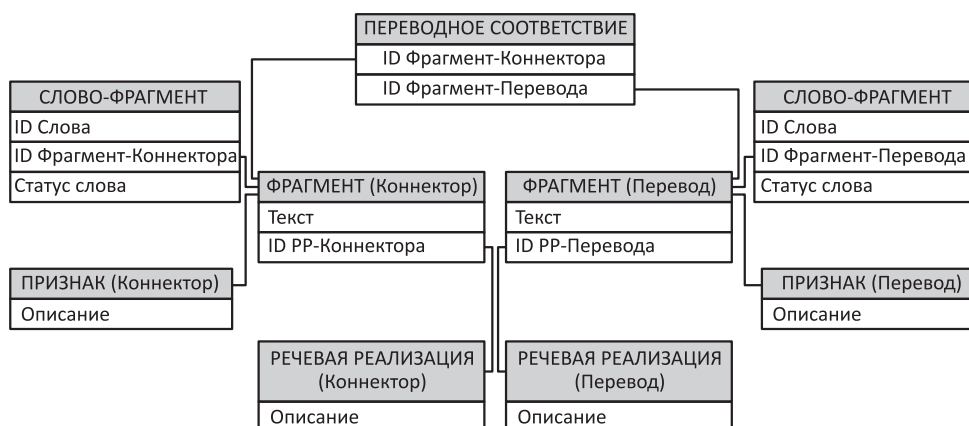


Рис. 1 Фрагмент логической структурной схемы НБД коннекторов, отвечающий за аннотацию переводных соответствий

ном контексте. Кроме того, для каждого фрагмента указано, какой именно РР он соответствует: для фрагмента оригинала это русский коннектор (в том виде, в каком он фигурирует в тексте), для фрагмента перевода — та часть французского фрагмента, которая, по мнению эксперта, в данном контексте считается его переводным эквивалентом (часто, но не всегда, это также коннектор). Наконец, каждому аннотируемому фрагменту присваивается ряд признаков (тегов), которые описывают особенности употребления данной РР в контексте данного фрагмента.

В табл. 2 приводится пример аннотации переводного соответствия, размеченного по описанной выше схеме.

Данное переводное соответствие построено для русского коннектора *a то*. В тексте фрагмента оригинала слова *a* и *то* выделены полужирным шрифтом,

Таблица 2 Пример аннотации переводного соответствия для русского коннектора *a то*

Фрагменты оригинала и перевода	Аннотации фрагментов
Чтобы Николка не вздумал тащить его. A то на улице... Слышишь?	Et que Nikolka ne s'avise pas de le porter. Sinon , dans la rue... Tu comprends?
РР: a то <ЛСО: причина> <предикация: без предикации> <позиция: начальная> <порядок: р CNT q> <статус: коннектор> <расположение: контакт>	РР: sinon <ЛСО: отрицательная альтернатива> <предикация: без предикации> <позиция: начальная> <порядок: р CNT q> <статус: коннектор>

что указывает на то, что они являются составными частями анализируемой РР. Остальной текст является контекстом, который помогает понять особенности употребления данной РР. В качестве признаков для фрагмента оригинала указаны:

- ЛСО, выражаемое данным коннектором (причина);
- отсутствие предикации во вводимом коннектором фрагменте текста;
- позиция коннектора в этом фрагменте (начальная);
- порядок следования фрагментов текста, соединяемых коннектором: сначала первый фрагмент (р), затем коннектор, затем второй фрагмент (q);
- статус РР — в данном случае она выполняет функцию коннектора (некоторые РР могут в одних контекстах быть коннекторами, а в других контекстах ими не быть);
- расположение элементов для неоднословных коннекторов (в данном случае элементы расположены контактно, но для некоторых коннекторов возможно дистантное расположение).

Структура и возможности НБД подробнее описаны в работах [15, 16].

4 Описание структуры неоднословных коннекторов

Предлагаемый информационно-компьютерный метод аннотирования неоднословных коннекторов в параллельных текстах предусматривает использование двухуровневой фасетной классификации: аннотированию подлежат, с одной стороны, конкретные употребления коннекторов в текстах корпуса (контекстное аннотирование), с другой стороны, аннотируется внутренняя структура каждого коннектора независимо от контекста его употребления (структурное аннотирование).

Структурная схема двухуровневого аннотирования РР-коннекторов показана на рис. 2. Употреблению коннектора соответствует таблица «Фрагмент (Коннектор)», при этом данная таблица связана с таблицей «Речевая реализация (Коннектор)», которая содержит различные зарегистрированные в НБД формы коннекторов.

Контекстное аннотирование осуществляется с помощью таблицы «Признак (Коннектор)». Как было показано в разд. 3, при контекстном аннотировании употребления коннекторов присваиваются признаки из ряда кластеров: выражаемое коннектором ЛСО, позиция коннектора в предложении, порядок следования вводимых коннектором фрагментов текста и т. д. При этом каждый признак входит ровно в один кластер, таким образом, классификация употреблений коннекторов представляет собой фасетную классификацию по заданному набору оснований.

В рамках структурного аннотирования коннекторы описываются внеконтекстно, в результате чего структурное описание однажды введенного в систему



Рис. 2 Фрагмент логической структурной схемы НБД коннекторов, отвечающий за аннотирование коннекторов

коннектора ассоциируется со всеми описаниями употреблений этого коннектора, что позволяет существенно сэкономить время разметчиков и избежать расхождений в описаниях. Структурное аннотирование коннекторов осуществляется с помощью таблицы «Кластер РР» — для каждой новой РР в НБД определяются кластеры, в которые она входит, причем одна и та же РР может одновременно входить в несколько кластеров.

В настоящее время структурное аннотирование проводится по двум основаниям. Во-первых, для каждого коннектора указывается, к какому структурному типу он принадлежит, и, во-вторых, для неоднословных коннекторов выделяются их структурные составляющие.

Выделяются следующие структурные типы коннекторов:

- одноэлементные — состоят из одного элемента (например, *и, или, но, а, хотя, иначе* и др.);
- многоэлементные — состоят из нескольких элементов (например, *и|вообще, к тому же, а особенно, но при всем при том* и др.);
- двухкомпонентные — коннекторы, части которых вводят два разных текстовых фрагмента (например, *даже если бы||то, как только||тут же, не только||а прежде всего, раз уж||так уж* и др.)¹;
- многокомпонентные — коннекторы, части которых вводят более двух разных текстовых фрагментов (например, *не только||но даже и||и даже, хотя||хотя||однако все же* и др.).

¹ Двойная вертикальная черта «||» в форме РР указывает на то, что данный коннектор состоит из нескольких компонентов, а одиночная вертикальная черта «|» — что языковые единицы, составляющие РР или ее компонент, разделены текстом.

Таблица 3 Примеры структурной аннотации РР-коннекторов с указанием кластеров, в которые они входят

		Речевые реализации		
		<i>а</i> <i>впрочем</i>	<i>да</i> <i>еще</i> <i>вдобавок</i>	<i>не только</i> <i>но</i> <i>также</i>
Кластеры РР	«Многоэлементные»	«Многоэлементные»	«Двухкомпонентные»	
	«а»	«вдобавок»	«не»	
	«впрочем»	«да»	«не только»	
		«еще»	«но»	
			«также»	
		«только»		

Следует заметить, что на 20.08.2018 в НБД было зарегистрировано 1276 русских РР-коннекторов, чуть более 5% из которых являются одноэлементными, 55% — многоэлементными, 34% — двухкомпонентными и 5% — многокомпонентными.

Кроме того, для неоднословных коннекторов указываются их значимые лексические составляющие (*не*, *то*, *более* и др.; некоторые из них могут самостоятельно функционировать в качестве коннекторов, например *а* или *и*). Возможность приписывать одну и ту же РР к разным кластерам позволяет устранить на этапе пополнения НБД такие трудноразрешимые проблемы, как определение соответствующего «базового» вида для различных форм коннекторов. Например, коннектор *но* и будет одновременно входить в кластер «но» и в кластер «и», коннектор *а также* будет одновременно входить в кластер «а» и кластер «также» и т. д. В табл. 3 приводятся три примера структурного аннотирования РР-коннекторов.

Описанная выше система «перекрестных» кластеров позволяет задавать комбинированные кластеры единиц, которые исследователи могут анализировать с помощью поисковых и статистических функций, реализованных в НБД (см., например, [17]). Так, задав поиск совместно по кластерам «но» и «даже», пользователь получит комбинированный кластер, включающий в себя 18 разных форм коннекторов: *но даже*, *но даже и*, *но даже просто*, *не только||но даже*, *хотя и||но даже и*, *не только||но и||и даже* и т. д. (рис. 3).

Использование двухуровневого подхода к описанию структуры неоднословных коннекторов позволяет избежать, с одной стороны, субъективности в присвоении статуса языковой единицы некоторым сочетаниям, оставляя за бортом другие, имеющие тот же статус (см. постановку вопроса в [4]), с другой — применения морфологической разметки, ориентирующейся на орфографическое слово, т. е. отделяемое пробелом. Наконец, применяемая в НБД система аннотирования позволяет дать более точное представление о системе связующих средств русского языка, их составе и частотности употребления (см. пример такого описания в [18]).

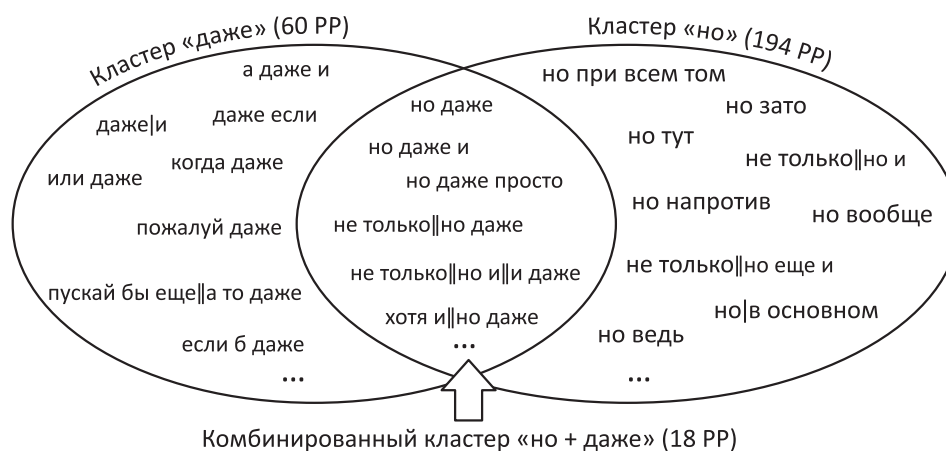


Рис. 3 Пример формирования комбинированного кластера

5 Заключение

Предлагаемый метод описания структуры неоднословных коннекторов опирается на когнитивно-семантический подход, в рамках которого предполагается, что неоднословные коннекторы являются относительно свободными сочетаниями, порождаемыми говорящим в зависимости не только от знания языка, но и от цели высказывания. В НБД для их аннотации используется двухуровневый подход: с одной стороны, аннотируются употребления коннекторов, с другой стороны, осуществляется структурная аннотация самих форм (PP) этих коннекторов. Для PP указывается, к какому структурному типу они принадлежат (одно- и много-элементные, двух- и n -компонентные) и какие структурные элементы они в себя включают.

Система перекрестных кластеров, используемая для аннотации неоднословных коннекторов, позволяет фиксировать сочетаемость единиц, служащих их составными частями, устраняя необходимость в распределении различных форм коннекторов по фиксированным независимым «гнездам», что сделать порой непросто, если вообще возможно. Возможность совместно задавать при поиске несколько кластеров позволяет конструировать комбинированные кластеры, которые можно далее анализировать с помощью поисковых и статистических функций, доступных в НБД.

Собранная в НБД информация о языковых единицах русского языка, выполняющих связующую функцию, уже позволяет существенно дополнить — увеличив его почти в два раза — список РГ-80 (1276 зарегистрированных в НБД PP против 666 языковых единиц в РГ-80), а также заполнить многие «белые пятна» в таблице сочетаемостных возможностей коннекторов, данной в работе [8].

Литература

1. *Инькова-Манзотти О. Ю.* Коннекторы противопоставления во французском и русском языках. Сопоставительное исследование. — М.: Информэлектро, 2001. 429 с.
2. Русская грамматика. Т. 2: Синтаксис / Под ред. Н. Ю. Шведовой. — М.: Наука, 1980. 714 с.
3. *Инькова О. Ю., Попкова Н. А.* Структура двухместных коннекторов русского языка в свете корпусных данных // Компьютерная лингвистика и интеллектуальные технологии: по мат-лам ежегодной Междунар. конф. «Диалог». — М.: РГГУ, 2016. Вып. 15(22). С. 763–775.
4. *Инькова О. Ю.* К проблеме описания многокомпонентных коннекторов русского языка: *не только... но и* // Вопросы языкознания, 2016. № 2. С. 37–60.
5. *Амеличева В. М.* Формальное и семантическое варьирование русского коннектора *не только... но и* и его французские эквиваленты // Съпоставително езикознание, 2017. Т. 42. № 4. С. 9–20.
6. *Кружков М. Г.* Информационные ресурсы контрастивных лингвистических исследований: электронные корпуса текстов // Системы и средства информатики, 2015. Т. 25. № 2. С. 140–159.
7. *Мустайоки А., Копотев М.* К вопросу о статусе эквивалентов слова типа *потому что, в зависимости от, к сожалению* // Вопросы языкознания, 2004. № 3. С. 88–107.
8. *Черемисина М. И., Колосова Т. А.* Очерки по теории сложного предложения. — 2-е изд. — М.: УРСС, 2010. 226 с.
9. *Danlos L., Colinet M., Steinlin J.* FDTB1, première étape du projet “French Discourse Treebank.” Repérage des connecteurs de discours en corpus // Discours, 2015. No. 17. <http://discours.revues.org/9065>.
10. *Knott A.* A data-driven methodology for motivating a set of coherence relations. — Edinburgh: University of Edinburgh, 1996. Ph.D. Thesis. <https://www.era.lib.ed.ac.uk/handle/1842/583>.
11. *Marcu D.* The rhetorical parsing, summarization and generation of natural language texts. — Toronto: University of Toronto, 1997. Ph.D. Thesis. <http://ftp.cs.toronto.edu/pub/gh/Marcu-PhDthesis.pdf>.
12. *Alonso L., Castellon I., Padro L.* Lexicón computacional de marcadores del discurso // Procesamiento Language Natural, 2002. No. 29. P. 239–246.
13. *Roze Ch., Danlos L., Muller Ph.* LEXCONN: A French lexicon of discourse connectives // Discours, 2012. No. 10. <http://discours.revues.org/8645>.
14. *Инькова О. Ю.* Надкорпусная база данных как инструмент формальной вариативности коннекторов // Компьютерная лингвистика и интеллектуальные технологии: по мат-лам ежегодной Междунар. конф. «Диалог». — М.: РГГУ, 2018. Вып. 17(24). С. 240–253.
15. *Kruzhkov M.* Supracorpora databases as corpus-based superstructure for manual annotation of parallel corpora // 8th Conference (International) on Corpus Linguistics / Eds. A. M. Ortiz, C. Pérez-Hernández. — EPiC ser. in language and linguistics, 2016. Vol. 1. P. 236–248. <https://easychair.org/publications/paper/jFjs>.
16. *Зацман И. М., Инькова О. Ю., Кружков М. Г., Попкова Н. А.* Представление кроссызыковых знаний о коннекторах в надкорпусных базах данных // Информатика и её применения, 2016. Т. 10. Вып. 1. С. 106–118.

17. Зацман И. М., Кружков М. Г., Лоцилова Е. Ю. Методы анализа частотности моделей перевода коннекторов и обратимость генерализации статистических данных // Системы и средства информатики, 2017. Т. 27. № 4. С. 164–176.
18. Кобозева И. М. Коннекторы контактного предшествования во французском и русском языках по данным параллельного корпуса // Съпоставително езикознание, 2017. Т. 42. № 4. С. 48–62.

Поступила в редакцию 05.09.18

METHOD FOR DESCRIPTION OF MULTIWORD CONNECTIVES IN SUPRACORPORA DATABASES

O. Yu. Inkova and M. G. Kruzikov

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

Abstract: This article presents a new method for describing the structure of multiword connectives implemented in the Supracorpora database (SCDB) of connectives. Currently, the structure of connectives is underinvestigated, and criteria for determining boundaries of connectives and their components are lacking. The proposed method is based on the cognitive-semantic approach that considers multiword connectives as more or less free word combinations generated in the process of speech. A two-tier faceted classification is proposed which allows annotating, on one hand, specific tokens of connectives in texts (context annotation) and, on the other hand, the inner structure of connectives (structural annotation). The structural annotation is based on two aspects: structural type and structural components of connectives. Based on the proposed annotation method, a system of cross-clusters is implemented that extends the search and statistical capabilities of SCDB. In addition, this method allows researchers to eliminate subjectivity during the annotation process and to fill some gaps in linguistic knowledge, for example, to gather new data on combinatorial capabilities of Russian connectives.

Keywords: connectives; linguistic items structure; linguistic items variation; corpus linguistics; annotation; faceted classification; supracorpora databases

DOI: 10.14357/08696527180416

Acknowledgments

This research was performed in the Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, with financial support of the Russian Foundation for Basic Research (grant No. 16-06-00070).

References

1. In'kova-Manzotti, O. Yu. 2001. *Konnektory protivopostavleniya vo frantsuzskom i russkom yazykakh. Sopostavitel'noe issledovanie* [Connectives of opposition in French and Russian. Comparative research]. Moscow: Informelektro. 429 p.
2. Shvedova, N. Yu., ed. 1980. *Russkaya grammatika. T. 2. Sintaksis* [Russian grammar. Vol. 2. Syntax]. Moscow: Nauka. 714 p.
3. In'kova, O. Yu., and N. A. Popkova. 2016. Struktura dvukhmestnykh konnektorov russkogo yazyka v svete korpusnykh dannykh [Structure of two-part Russian connectives based on corpus data]. *Computer Linguistic and Intellectual Technologies: Conference (International) "Dialog" Proceedings*. Moscow: RGGU. 15(22):763–775.
4. Inkova, O. Yu. 2016. K probleme opisaniya mnogokomponentnykh konnektorov russkogo yazyka: *ne tol'ko... no i* [Towards the description of multiword connectives in Russian: *ne tol'ko... no i* 'not only... but also']. *Voprosy yazykoznaniya* [Topics in the Study of Language] 2:37–60.
5. Amelicheva, V. M. 2017. Formal'noe i semanticheskoe var'irovanie russkogo konnektora *ne tol'ko... no i* i ego frantsuzskie ekvivalenty [Formal and semantic variation of the Russian connective *ne tol'ko... no i* (*not only... but also*) and its French equivalents]. *Contrastive Linguistics* 42(4):9–20.
6. Kruzhkov, M. G. 2015. Informatsionnye resursy kontrastivnykh lingvisticheskikh issledovaniy: elektronnye korpusa tekstov [Information resources for contrastive studies: Electronic text corpora]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 25(2):140–159.
7. Mustayoki, A., and M. Kopotev. 2004. K voprosu o statuse ekvivalentov slova tipa *potomu chto, v zavisimosti ot, k sozhaleniyu* [On the status of such word equivalents as *potomu chto, v zavisimosti ot, k sozhaleniyu* (because, depending on, unfortunately)]. *Voprosy yazykoznaniya* [Topics in the Study of Language] 3:88–107.
8. Cheremisina, M. I., and T. A. Kolosova. 2010. *Ocherki po teorii slozhnogo predlozheniya* [Studies on the theory of the complex sentence]. 2nd ed. Moscow: URSS. 226 p.
9. Danlos, L., M. Colinet, and J. Steinlin. 2015. FDTB1, première étape du projet "French Discourse Treebank:" Repérage des connecteurs de discours en corpus. *Discours* 17. Available at: <http://discours.revues.org/9065> (accessed August 29, 2018).
10. Knott, A. 1996. A data-driven methodology for motivating a set of coherence relations. Edinburg: University of Edinburgh. Ph.D. Thesis. Available at: <https://www.era.lib.ed.ac.uk/handle/1842/583> (accessed August 29, 2018).
11. Marcu, D. 1997. The rhetorical parsing, summarization and generation of natural language texts. Toronto: University of Toronto. Ph.D. Thesis. Available at: <http://ftp.cs.toronto.edu/pub/gh/Marcu-PhDthesis.pdf> (accessed August 29, 2018).
12. Alonso, L., I. Castellon, and L. Padro. 2002. Lexicón computacional de marcadores del discurso. *Procesamiento Lenguaje Natural* 29:239–246.
13. Roze, Ch., L. Danlos, and Ph. Muller. 2012. LEXCONN: A French lexicon of discourse connectives. *Discours* 10. Available at: <http://discours.revues.org/8645> (accessed August 29, 2018).
14. In'kova, O. Yu. 2018. Nadkorpurnaya baza dannykh kak instrument formal'noy variativnosti konnektorov [Supracorpora database as an instrument for formal variation of

- connectives]. *Computer Linguistic and Intellectual Technologies: Conference (International) "Dialog" Proceedings*. Moscow: RGGU. 17(24):240–253.
15. Kruzhkov, M. 2016. Supracorpora databases as corpus-based superstructure for manual annotation of parallel corpora. *8th Conference (International) on Corpus Linguistics*. Eds. A. M. Ortiz and C. Pérez-Hernández. EPiC ser. in language and linguistics. 1:236–248. Available at: <https://easychair.org/publications/paper/jFjs> (accessed August 29, 2018).
 16. Zatsman, I. M., O. Yu. In'kova, M. G. Kruzhkov, and N. A. Popkova. 2016. Predstavlenie krossyazykovykh znaniy o konnektorakh v nadkorporusnykh bazakh dannykh [Representation of cross-lingual knowledge about connectors in supracorpora databases]. *Informatika i ee Primeneniya — Inform. Appl.* 10(1):106–118.
 17. Zatsman, I. M., M. G. Kruzhkov, and E. Yu. Loshchilova. 2017. Metody analiza chastotnosti modeley perevoda konnektorov i obratimost' generalizatsii statisticheskikh dannykh [Methods of frequency analysis of connectives translations and reversibility of statistical data generalization]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 27(4):164–176.
 18. Kobozeva, I. M. 2017. Konnektory kontaktnogo predshestvovaniya vo frantsuzskom i russkom yazykakh po dannym parallel'nogo korpusa [Connectives of immediate precedence in Russian and French based on parallel corpus data]. *Contrastive Linguistics* 42(4):48–62.

Received September 5, 2018

Contributors

Inkova Olga Yu. (b. 1965)— Doctor of Science (PhD) 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; olyainkova@yandex.ru

Kruzhkov Mikhail G. (b. 1975) — 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; magnit75@yandex.ru

ТЕМАТИЧЕСКИЕ ИНДИКАТОРЫ ПЕРЕНОСА НАУЧНЫХ ЗНАНИЙ В СФЕРУ ТЕХНОЛОГИЙ*

В. А. Минин¹, И. М. Зацман², В. А. Хавансков³, С. К. Шубников⁴

Аннотация: Индикаторное оценивание процесса переноса знаний является важным элементом в исследовании информационных взаимосвязей науки и процессов разработки новых технологий, характеризующим процессы создания научно-технического задела в области перспективных технологий. В процессе изучения информационных взаимосвязей проведен эксперимент по вычислению значений индикаторов переноса научных знаний в сферу разработки информационных технологий (ИТ). С этой целью проанализирован одиннадцатилетний массив полнотекстовых описаний изобретений, полученных из Роспатента. Сопоставлено смысловое содержание индикаторов переноса, характеризующих информационные взаимодействия различных направлений научных исследований (НИИ) и групп технологий. Порядок проведения эксперимента описан в статье авторов, опубликованной в предыдущем номере журнала. В данной статье, которая является ее продолжением, обсуждаются основные понятия, использованные в ходе эксперимента, и приведены новые результаты вычисления значений индикаторов.

Ключевые слова: информационные взаимосвязи науки и технологий; цитирование научных работ; интенсивность процессов переноса знаний; индикаторное оценивание; информационные технологии

DOI: 10.14357/08696527180417

1 Основные положения

Статья завершает цикл исследований и публикаций по проекту РФФИ «Методы и информационные технологии оценивания интенсивности переноса научных знаний в сферу разработки технологий», ориентированному на развитие методологии оценки инновационного потенциала НИИ [1–5].

* Исследование выполнено в Институте проблем информатики Федерального исследовательского центра «Информатика и управление» при финансовой поддержке РФФИ (проект 16-07-00075).

¹ Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, aleksiss@ya.ru

² Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, izatsman@yandex.ru

³ Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, chavanskov@yandex.ru

⁴ Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, sergeysh50@yandex.ru

Цель исследований — получить не интегральную (обобщенную) оценку интенсивности переноса знаний по всем областям знаний без учета их структурирования на научные дисциплины [2], а тематические индикаторные оценки на основе классификатора, описывающего то или иное деление всей совокупности знаний на НИИ. Для этого разработаны и используются тематические индикаторы, которые характеризуют различные составляющие обобщенной интенсивности переноса научных знаний из различных научных направлений в исследуемую группу технологий. Таким образом, тематические индикаторы дают представление о структуре потока научных знаний, переносимых в технологическую сферу.

В начале статьи остановимся на основных положениях методологии, создаваемой участниками проекта более 10 лет [6–12].

1. Информационно-методический подход (с применением индикаторов, характеризующих информационные аспекты переноса научных знаний), разработанный авторским коллективом и описанный в [2], позволяет зафиксировать в описаниях изобретений ссылки на публикации (цитирование), содержащие результаты научных исследований, которые рассматриваются как признак потенциального воздействия результатов фундаментальной и прикладной науки на создание патентуемых технологий. Подход основан на компьютерной обработке и многоаспектном анализе полнотекстовых описаний изобретений, патентуемых Роспатентом. При этом изобретение трактуется как новое технологическое решение.
2. Для рубрикации цитируемых публикаций применяются классификаторы научных направлений: «Классификатор РФФИ» (фундаментальная наука) и «Классификатор ГРНТИ» (фундаментальная и прикладная наука). В этих классификаторах научных направлений рубрики имеют названия и обозначаются кодами. Для описания анализируемых технологических областей используются индексы Международной патентной классификации (МПК), содержащей 5 уровней детализации описания технологий (разделы, классы, подклассы, группы и подгруппы технологий) [5].
3. В описаниях изобретений присутствуют ссылки на научные публикации, содержащие информацию (как форму представления научного знания) о результатах, полученных в ходе научных исследований. Публикации, цитируемые в патентуемом изобретении, позволяют после их рубрикации выявить информационные связи между различными НИИ и группами технологий. Такие связи могут свидетельствовать об использовании результатов этих исследований. Сам факт цитирования в изобретении научных публикаций по какому-либо НИИ может говорить о потенциальном воздействии данного НИИ на ту группу технологий, к которой относится это изобретение. Чем чаще результаты данного НИИ цитируются, тем, скорее всего, значимее степень его воздействия на развитие анализируемой группы технологий.

4. Интенсивность воздействия различных ННИ на группы технологий оценивается с помощью количественных индикаторов интенсивности переноса (ИИП) научных знаний, которые характеризуют частоту цитирований публикаций с описанием научных результатов данного ННИ в изобретениях, относящихся к той или иной группе технологий.

Обзор и анализ зарубежного опыта по тематике проекта приведен в работах [2, 7, 8].

2 Эксперимент по вычислению значений индикаторов

Для проведения эксперимента по вычислению значений ИИП научных знаний был разработан макет аналитической информационной системы (АИС), описание которого дано в работе [4]. Коротко изложим основные положения эксперимента.

1. Вычислялись ИИП научных знаний для группы ИТ. Цель эксперимента — сопоставление и уточнение смыслового содержания (семантики) ИИП различных видов. Эксперимент проводился на основе анализа более 8 тыс. описаний отечественных и зарубежных изобретений, запатентованных по указанной группе технологий в Роспатенте в 2002–2012 гг. Формирование ИИП научных знаний требует автоматизированной обработки больших объемов слабоструктурированных полнотекстовых описаний изобретений и сопоставления метаданных патентной информации с метаданными научных публикаций (в журналах или материалах конференций), цитируемых в описаниях изобретений. С этой целью был разработан и использован макет АИС [4, 5].
2. Индикаторы используются для оценивания результатов влияния научной деятельности на сферу технологий, т. е. для оценки инновационного потенциала различных научных направлений науки.
3. В соответствии с классификацией переносимых знаний одна категория индикаторов отражает использование в изобретениях (технологиях) результатов фундаментальных исследований, другая — прикладных научных исследований.
4. Виды индикаторов внутри каждой категории выделяются в зависимости от способа учета рубрик классификаторов: (а) «фактический», когда учитывается только факт наличия какой-либо рубрики классификатора у одной или нескольких публикаций — без учета числа таких повторений в одном изобретении; (б) «частотный», когда учитывается число повторений рубрик классификаторов в одном изобретении [2]. Фактический способ учета позволяет, в частности, сопоставить интенсивности упоминаний ННИ в описании изобретений, относящихся к разным группам технологий, так как значения индикаторов этого вида нормализованы по средней частоте

повторений в группе, которая в разных группах может существенно отличаться. Частотный способ учета позволяет определить среднее значение частоты цитирования публикаций одного ННИ в пределах группы (для одного изобретения).

3 Индикаторы переноса научных знаний в сферу технологий

В работе [5] представлена в качестве примера полученная в ходе эксперимента подборка значений индикаторов фактического и частотного вида, которые характеризуют интенсивность переноса в ИТ знаний из 29 ННИ согласно рубрикам классификатора РФФИ (табл. 1).

Таблица 1 Значения ИИП для двух способов учета рубрик классификатора РФФИ

Классификатор РФФИ		Способ учета рубрик классификатора РФФИ			
Код рубрики	Название рубрики	Фактический учет		Частотный учет	
		Абсолютное значение ИИП	Доля рубрики, %	Абсолютное значение ИИП	Доля рубрики, %
01-200	ИНФОРМАТИКА	146	19,05	52,66	23,47
07-500	СИСТЕМЫ ОБРАБОТКИ И АНАЛИЗА ДАННЫХ	97	12,66	30,95	13,79
07-800	ФУНДАМЕНТАЛЬНЫЕ ОСНОВЫ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ И ТЕЛЕКОММУНИКАЦИЙ	89	11,61	19,14	8,54
07-700	ЭЛЕМЕНТНАЯ БАЗА ИНФОРМАЦИОННЫХ СИСТЕМ	33	4,30	18,49	8,24
08-600	ТЕХНИЧЕСКИЕ СИСТЕМЫ И ПРОЦЕССЫ УПРАВЛЕНИЯ	53	6,91	15,85	7,06
07-600	ВЫЧИСЛИТЕЛЬНЫЕ И СЕТЕВЫЕ РЕСУРСЫ	46	6,00	13,02	5,80
07-400	ИНФОРМАЦИОННЫЕ РЕСУРСЫ И СИСТЕМЫ	58	7,57	9,92	4,42
01-100	МАТЕМАТИКА	50	6,52	9,60	4,28
02-400	РАДИОФИЗИКА, ЭЛЕКТРОНИКА, АКУСТИКА	35	4,56	9,48	4,22
02-200	ФИЗИКА КОНДЕНСИРОВАННЫХ СРЕД	34	4,43	8,87	3,95
08-300	ЭЛЕКТРОФИЗИКА, ЭЛЕКТРОТЕХНИКА И ЭЛЕКТРОЭНЕРГЕТИКА	13	1,69	8,68	3,87

Окончание табл. 1 на с. 186

Таблица 1 (окончание) Значения ИИП для двух способов учета рубрик классификатора РФФИ

Код рубрики	Классификатор РФФИ Название рубрики	Способ учета рубрик классификатора РФФИ			
		Фактический учет		Частотный учет	
		Абсолютное значение ИИП	Доля рубрики, %	Абсолютное значение ИИП	Доля рубрики, %
02-300	ОПТИКА, КВАНТОВАЯ ЭЛЕКТРОНИКА	27	3,52	7,05	3,14
04-300	ФУНДАМЕНТАЛЬНАЯ МЕДИЦИНА И ФИЗИОЛОГИЯ	10	1,30	4,90	2,18
05-700	ГЕОГРАФИЯ И ГИДРОЛОГИЯ СУШИ	12	1,56	3,47	1,55
06-200	ЭКОНОМИЧЕСКИЕ НАУКИ	4	0,52	2,10	0,94
05-400	ГЕОФИЗИКА	6	0,78	1,97	0,88
06-300	ФИЛОСОФСКИЕ И СОЦИАЛЬНЫЕ НАУКИ, ПСИХОЛОГИЯ	12	1,56	1,60	0,71
04-200	ФИЗИКО-ХИМИЧЕСКАЯ БИОЛОГИЯ	6	0,78	1,52	0,68
02-100	ЯДЕРНАЯ ФИЗИКА	5	0,65	0,83	0,37
02-900	МЕДИЦИНСКАЯ ФИЗИКА	9	1,17	0,72	0,32
05-500	ОКЕАНОЛОГИЯ	2	0,26	0,62	0,27
08-100	МАШИНОВЕДЕНИЕ И ИНЖЕНЕРНАЯ МЕХАНИКА	3	0,39	0,61	0,27
02-600	ФИЗИКА ПЛАЗМЫ	2	0,26	0,57	0,25
03-200	НЕОРГАНИЧЕСКАЯ ХИМИЯ	1	0,13	0,40	0,18
04-100	ОБЩАЯ БИОЛОГИЯ	6	0,78	0,40	0,18
06-400	ФИЛОЛОГИЯ, КУЛЬТУРОЛОГИЯ, ИСКУССТВОВЕДЕНИЕ, ОХРАНА ПАМЯТНИКОВ ИСТОРИИ И КУЛЬТУРЫ	2	0,26	0,33	0,15
05-600	ФИЗИКА АТМОСФЕРЫ	2	0,26	0,31	0,14
02-800	АСТРОНОМИЯ	3	0,39	0,30	0,13
02-700	ТЕОРЕТИЧЕСКАЯ ФИЗИКА	1	0,13	0,06	0,02

На основе данных табл. 1 построена диаграмма, демонстрирующая соотношение фактических и частотных значений упоминаний в описаниях ИТ-изобретений 29 ННИ согласно рубрикам классификатора РФФИ (рис. 1).

Аналогичным образом построена диаграмма на рис. 2, демонстрирующая соотношение фактических и частотных значений упоминаний 40 ННИ в описаниях ИТ-изобретений согласно рубрикам классификатора ГРНТИ.

В целях дальнейшего анализа выделим для сопоставления по 10 наиболее часто встречающихся рубрик ННИ из классификаторов РФФИ и ГРНТИ, ко-

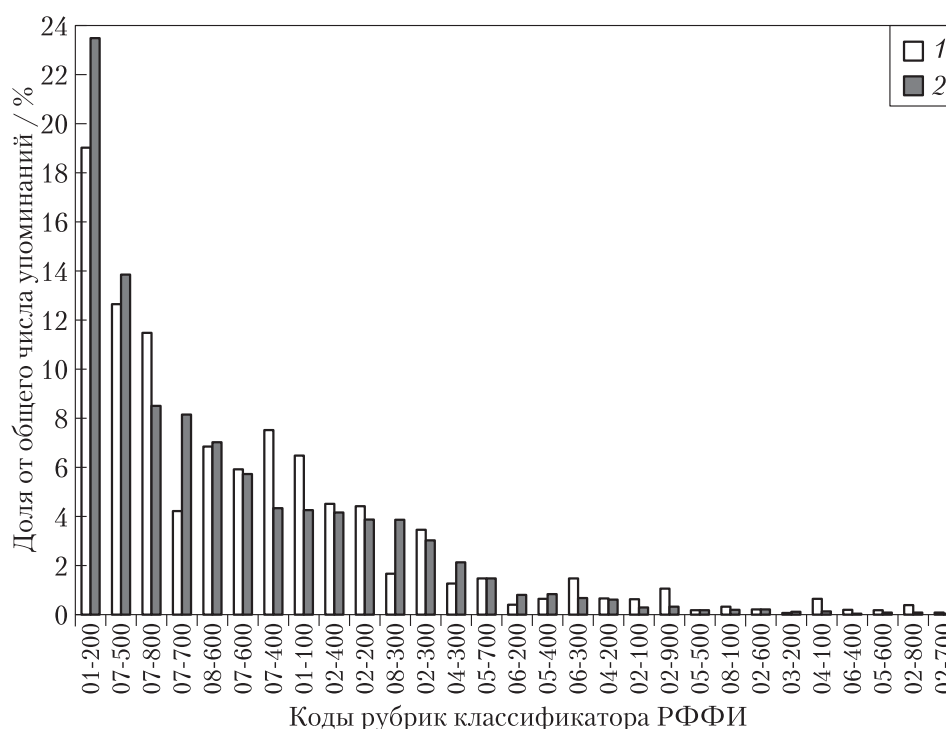


Рис. 1 Сравнение фактических и частотных значений упоминания 29 ННИ согласно классификатору РФФИ: 1 — фактические значения; 2 — частотные значения

торые приведены в табл. 2 и 3 соответственно. Эти таблицы дают представление о степени взаимосвязи ИТ с различными направлениями науки. Расположение ННИ в них отображает уровень воздействия на ИТ результатов научных исследований по соответствующей рубрике согласно вычисленным значениям ИИП научных знаний (по убыванию значений).

Сопоставление данных табл. 2 и 3 показывает тематическую структуру потоков знаний, переносимых в группу ИТ.

Отметим важный терминологический аспект в названиях рубрик классификаторов, отраженный ранее в работе [5]. Термин «Информатика» в классификаторах РФФИ и ГРНТИ обозначает две разные научные дисциплины. В РФФИ это компьютерная наука (computer science), а в ГРНТИ — информационная наука (information science), которая на современном этапе является профильной, но не самой частотной научной дисциплиной с точки зрения создания ИТ в период 2002–2012 гг.

Объединяя табл. 2 и 3, переходим к табл. 4, которая дает комплексное представление о спектре использования в группе ИТ-изобретений результатов

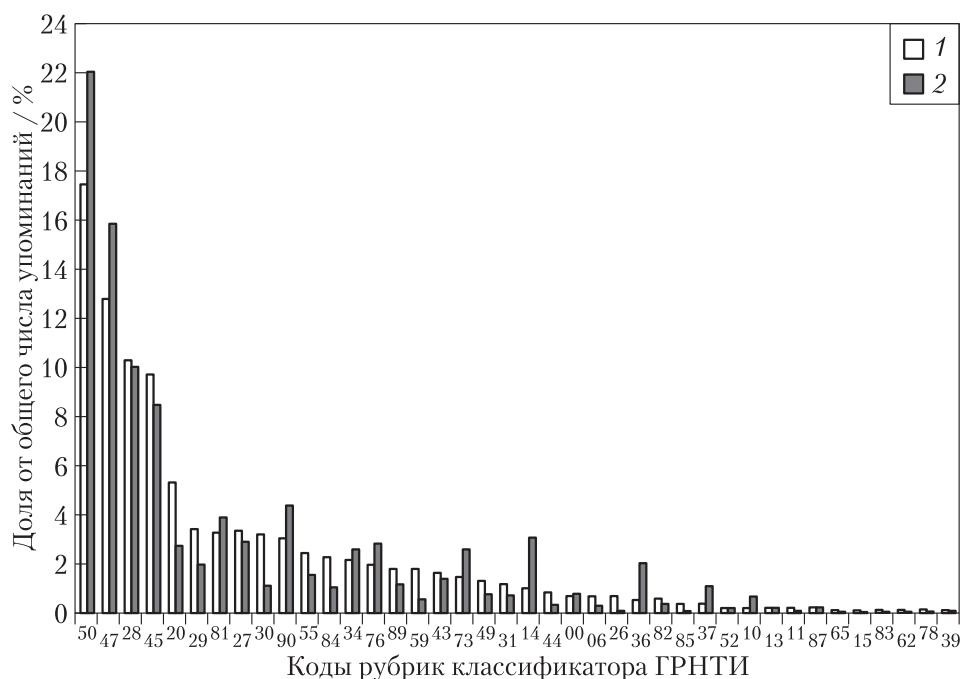


Рис. 2 Сравнение фактических и частотных значений упоминания 40 ННИ согласно ГРТИ: 1 — фактические значения; 2 — частотные значения

Таблица 2 Список 10 рубрик классификатора РФФИ, наиболее часто упоминаемых в ИТ-изобретениях

Код рубрики классификатора РФФИ	Название рубрики
01-200	ИНФОРМАТИКА
07-500	СИСТЕМЫ ОБРАБОТКИ И АНАЛИЗА ДАННЫХ
07-800	ФУНДАМЕНТАЛЬНЫЕ ОСНОВЫ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ И ТЕЛЕКОММУНИКАЦИЙ
07-700	ЭЛЕМЕНТНАЯ БАЗА ИНФОРМАЦИОННЫХ СИСТЕМ
08-600	ТЕХНИЧЕСКИЕ СИСТЕМЫ И ПРОЦЕССЫ УПРАВЛЕНИЯ
07-600	ВЫЧИСЛИТЕЛЬНЫЕ И СЕТЕВЫЕ РЕСУРСЫ
07-400	ИНФОРМАЦИОННЫЕ РЕСУРСЫ И СИСТЕМЫ
01-100	МАТЕМАТИКА
02-400	РАДИОФИЗИКА, ЭЛЕКТРОНИКА, АКУСТИКА
02-200	ФИЗИКА КОНДЕНСИРОВАННЫХ СРЕД

Таблица 3 Список 10 рубрик классификатора ГРНТИ, наиболее часто упоминаемых в ИТ-изобретениях

Код рубрики классификатора ГРНТИ	Название рубрики
50	АВТОМАТИКА. ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА
47	ЭЛЕКТРОНИКА. РАДИОТЕХНИКА
28	КИБЕРНЕТИКА
45	ЭЛЕКТРОТЕХНИКА
20	ИНФОРМАТИКА
29	ФИЗИКА
81	ОБЩИЕ И КОМПЛЕКСНЫЕ ПРОБЛЕМЫ ТЕХНИЧЕСКИХ И ПРИКЛАДНЫХ НАУК И ОТРАСЛЕЙ НАРОДНОГО ХОЗЯЙСТВА
27	МАТЕМАТИКА
30	МЕХАНИКА
90	МЕТРОЛОГИЯ

исследований по различным направлениям науки, что, собственно, и характеризует тематическое воздействие науки на данную группу технологий.

В табл. 4 рубрики ННИ классификаторов РФФИ и ГРНТИ упорядочены по степени их влияния на группу ИТ. Обратим внимание на изменение позиций рубрик классификаторов в зависимости от того, какой именно способ учета применяется — «фактический» или «частотный». Например, в табл. 4 рубрика «ИНФОРМАЦИОННЫЕ РЕСУРСЫ И СИСТЕМЫ» классификатора РФФИ при изменении способа учета перемещается с 4-й позиции (упорядочение по фактическому значению) на 7-ю (упорядочение по частотности).

Такое перемещение говорит о том, что фактически эта тематика цитируется во всей группе ИТ-изобретений чаще, чем следующие за ней, включая три тематики под номерами 5, 6 и 7, но по частотности цитирования в рамках *одного изобретения* эта тематика цитируется реже, чем предшествующие ей. Отметим, что тематика под номером 6 («МАТЕМАТИКА») стала номером 8, а тематика «ВЫЧИСЛИТЕЛЬНЫЕ И СЕТЕВЫЕ РЕСУРСЫ» переходит с позиции 7 на позицию 6. Аналогичные изменения можно наблюдать и по другим рубрикам ННИ как для классификатора РФФИ, так и для ГРНТИ.

Приведенные результаты эксперимента относятся к рассмотрению взаимосвязи науки и технологий, так сказать, «со стороны технологий». Иными словами, показано, какие научные направления влияют на группу ИТ-изобретений. Но аналогичным образом можно посмотреть на взаимосвязь науки и технологий «со стороны научных исследований», т. е. определить, на создание каких ИТ оказали влияние результаты того или иного ННИ в период 2002–2012 гг.

Таблица 4 Рубрики ННИ, упорядоченные по фактическим и частотным их упоминаниям в ИТ-изобретениях

№	Упорядочение по фактическим значениям упоминания рубрик ННИ		Упорядочение по частотным значениям упоминания рубрик ННИ	
	Название рубрик РФФИ	Название рубрик ГРНТИ	Название рубрик РФФИ	Название рубрик ГРНТИ
1.	ИНФОРМАТИКА	АВТОМАТИКА. ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА	ИНФОРМАТИКА	АВТОМАТИКА. ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА
2.	СИСТЕМЫ ОБРАБОТКИ И АНАЛИЗА ДАННЫХ	ЭЛЕКТРОНИКА. РАДИОТЕХНИКА	СИСТЕМЫ ОБРАБОТКИ И АНАЛИЗА ДАННЫХ	ЭЛЕКТРОНИКА. РАДИОТЕХНИКА
3.	ФУНДАМЕНТАЛЬНЫЕ ОСНОВЫ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ И ТЕЛЕКОММУНИКАЦИЙ	КИБЕРНЕТИКА	ФУНДАМЕНТАЛЬНЫЕ ОСНОВЫ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ И ТЕЛЕКОММУНИКАЦИЙ	КИБЕРНЕТИКА
4.	ИНФОРМАЦИОННЫЕ РЕСУРСЫ И СИСТЕМЫ	ЭЛЕКТРОТЕХНИКА	ЭЛЕМЕНТНАЯ БАЗА ИНФОРМАЦИОННЫХ СИСТЕМ	ЭЛЕКТРОТЕХНИКА
5.	ТЕХНИЧЕСКИЕ СИСТЕМЫ И ПРОЦЕССЫ УПРАВЛЕНИЯ	ИНФОРМАТИКА как information science	ТЕХНИЧЕСКИЕ СИСТЕМЫ И ПРОЦЕССЫ УПРАВЛЕНИЯ	МЕТРОЛОГИЯ

6.	МАТЕМАТИКА	ФИЗИКА	ВЫЧИСЛИТЕЛЬНЫЕ И СЕТЕВЫЕ РЕСУРСЫ	ОБЩИЕ И КОМПЛЕКСНЫЕ ПРОБЛЕМЫ ТЕХНИЧЕСКИХ И ПРИКЛАДНЫХ НАУК И ОТРАСЛЕЙ НАРОДНОГО ХОЗЯЙСТВА
7.	ВЫЧИСЛИТЕЛЬНЫЕ И СЕТЕВЫЕ РЕСУРСЫ	ОБЩИЕ И КОМПЛЕКСНЫЕ ПРОБЛЕМЫ ТЕХНИЧЕСКИХ И ПРИКЛАДНЫХ НАУК И ОТРАСЛЕЙ НАРОДНОГО ХОЗЯЙСТВА	ИНФОРМАЦИОННЫЕ РЕСУРСЫ И СИСТЕМЫ	НАРОДНОЕ ОБРАЗОВАНИЕ. ПЕДАГОГИКА
8.	РАДИОФИЗИКА, ЭЛЕКТРОНИКА, АКУСТИКА	МАТЕМАТИКА	МАТЕМАТИКА	МАТЕМАТИКА
9.	ФИЗИКА КОНДЕНСИРОВАННЫХ СРЕД	МЕХАНИКА	РАДИОФИЗИКА, ЭЛЕКТРОНИКА, АКУСТИКА	МЕДИЦИНА И ЗДРАВООХРАНЕНИЕ
10.	ЭЛЕМЕНТНАЯ БАЗА ИНФОРМАЦИОННЫХ СИСТЕМ	МЕТРОЛОГИЯ	ФИЗИКА КОНДЕНСИРОВАННЫХ СРЕД	ИНФОРМАТИКА как information science

В качестве примера приведем полученный из описаний ИТ-изобретений перечень тех ИТ, представленных индексами подклассов МПК, на которые влияют результаты научных исследований только одного ННИ «Информатика» согласно классификатору РФФИ (табл. 5).

Таблица 5 Перечень технологий, представленных индексами подклассов МПК, на которые влияют результаты научных исследований по одному ННИ «Информатика» (классификатор РФФИ)

Индексы подклассов МПК	Название подклассов МПК
G06F	Обработка цифровых данных с помощью электрических устройств
G06K	Распознавание, представление и воспроизведение данных; манипулирование носителями информации; носители информации
G06T	Обработка или генерация данных изображения, вообще
G05B	Регулирующие и управляющие системы общего назначения; функциональные элементы таких систем; устройства для контроля или испытания таких систем или элементов
G06G	Аналоговые вычислительные машины
G06N	Компьютерные системы, основанные на специфических вычислительных моделях
G01S	Радиопеленгация; радионавигация; измерение расстояния или скорости с использованием радиоволн; определение местоположения или обнаружение объектов с использованием отражения или переизлучения радиоволн; аналогичные системы с использованием других видов волн
G06Q	Системы обработки данных или способы, специально предназначенные для административных, коммерческих, финансовых, управленческих, надзорных или прогностических целей; системы или способы, специально предназначенные для административных, коммерческих, финансовых, управленческих, надзорных или прогностических целей, не предусмотренные в других подклассах
G10L	Анализирование или синтезирование речи; распознавание речи
G08G	Системы регулирования движения транспортных средств
G09C	Шифровальные или дешифровальные устройства для тайнописи или других целей, связанных с секретной перепиской
G05F	Системы регулирования электрических или магнитных величин
G06E	Оптические вычислительные устройства
G01D	Измерения, специально не предназначенные для особых переменных величин; устройства или приборы для измерения двух или более переменных величин, не отнесенные к какому-либо одному подклассу; тарифные счетчики; измерения или испытания, не отнесенные к другим подклассам
G01R	Измерение электрических и магнитных величин

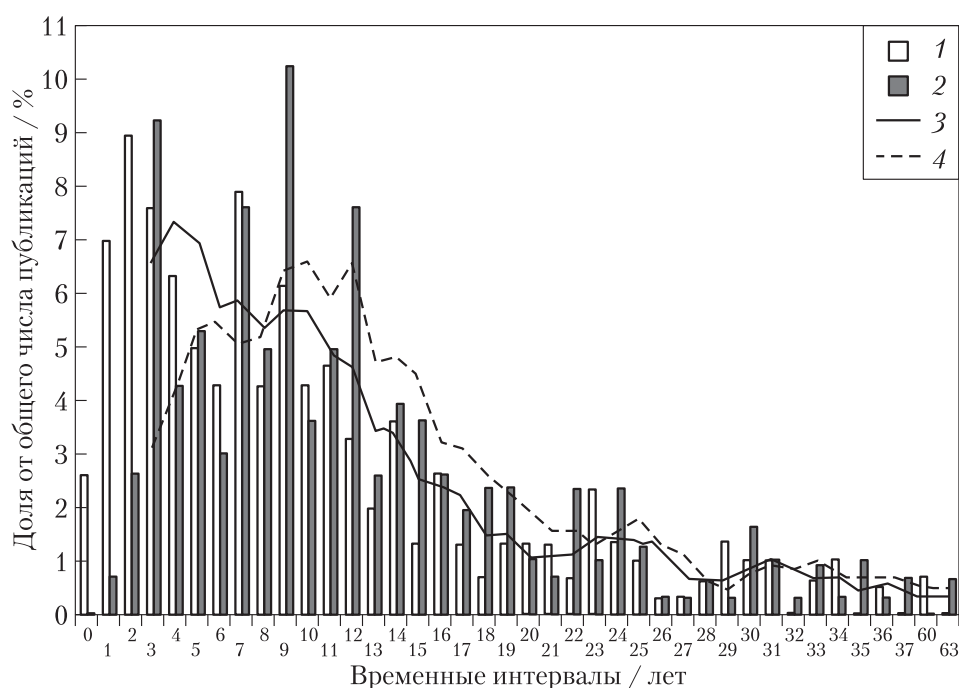


Рис. 3 Распределение времени отклика изобретений на публикацию научных результатов: 1 — относительно даты подачи заявки на изобретение; 2 — относительно даты публикации решения о выдаче патента на изобретение; 3 — линия тренда относительно даты подачи заявки; 4 — линия тренда относительно даты публикации

В заключение этого раздела рассмотрим вопрос, который также представляет значительный интерес — сколько долго опубликованные научные результаты сохраняют свою актуальность для процессов создания технологий? Долговременность воздействия научных результатов исследований по различным ННИ на группу ИТ-изобретений представлена на рис. 3, построенном с использованием данных, полученных в ходе эксперимента. Как видно из данной диаграммы, используя предложенный информационно-компьютерный подход, можно прогнозировать жизненный цикл научного результата с точки зрения процессов создания новых технологических решений.

Данная диаграмма иллюстрирует тот факт, что, несмотря на высокую динамику создания и развития ИТ, используемые результаты *научных исследований* сохраняют свою значимость (выявленную через цитируемые научные публикации) на протяжении достаточно больших отрезков времени, измеряемых десятками лет. Понижение линий трендов естественным образом объясняется тем, что с годами поток публикаций, в которых упоминаются полученные ранее научные результаты, падает.

Кроме того, на долговременность воздействия научного результата на технические решения иногда влияет отсутствие технологической базы для его промышленной реализации. В качестве примера обратим внимание на время получения первых научных результатов, оказавших влияние на создание LCD-мониторов¹. Жидкие кристаллы как состояние вещества были обнаружены в 1880-х гг. Фридрихом Рейнитцером (1858–1927), но лишь в наши дни они нашли массовое и многообразное прикладное применение [13]. Джон Л. Дженнинг в 1968 г. изобрел жидкокристаллические дисплеи (LCD), и уже в 1972 г. первый жидкокристаллический дисплей на активной матрице был создан в Соединенных Штатах Т. Питером Броуди и его сотрудниками [14].

4 Заключение

Результаты эксперимента подтверждают сделанный ранее вывод [5] о том, что разработанные в ходе выполнения проекта РФФИ 16-07-00075 и использованные в данном эксперименте методика, модель, а также концепция и архитектура макета информационной системы, как и проектные решения функциональных подсистем, дают возможность вычислять значения ИИП знаний из различных ННИ. Это позволяет экспертам оценить количественно степень воздействия различных ННИ на развитие той или иной группы технологий.

Разработанная совокупность тематических ИИП, разделенная по признакам фактического и частотного учета рубрик ННИ, отображает многоаспектность связей «наука–технологии». Анализ полученных результатов позволяет сделать вывод о том, что список тематических ИИП переноса знаний, упорядоченных по частотности упоминаний ННИ, является наглядной иллюстрацией степени воздействия того или иного ННИ на исследуемую технологическую сферу (в данном случае — ИТ).

В заключение отметим, что разработанная в проекте РФФИ 16-07-00075 совокупность индикаторов предназначена для использования *в сочетании* с экспертным оцениванием вычисленных значений ИИП научных знаний. Эта совокупность может быть использована для экспертного выявления тех поисковых исследований, вложение средств в которые может дать синергетический эффект в развитии отечественной технологической сферы.

Литература

1. Минин В. А., Зацман И. М., Хавансков В. А., Шубников С. К. Интенсивность цитирования научных публикаций в изобретениях по информационно-компьютерным технологиям, патентуемых в России отечественными и зарубежными заявителями // Информатика и её применения, 2016. Т. 10. Вып. 2. С. 107–122.

¹ Данные сведения любезно предоставлены В. В. Косариком — главным специалистом Института проблем информатики ФИЦ ИУ РАН.

2. Минин В. А., Зацман И. М., Хавансков В. А., Шубников С. К. Методы индикаторного оценивания процессов переноса знаний из области научных исследований в сферу технологического развития // Научно-техническая информация. Сер. 2, 2017. № 8. С. 1–12.
3. Zatsman I. M., Lukyanov G. V., Minin V. A., Havanskov V. A., Shubnikov S. K. Indicator evaluation of processes of knowledge transfer from science to technology // Информатика и её применения, 2017. Т. 11. Вып. 3. С. 132–141.
4. Минин В. А., Зацман И. М., Хавансков В. А., Шубников С. К. Макетирование информационной системы вычисления индикаторов переноса научных знаний в сферу технологий // Системы и средства информатики, 2017. Т. 27. № 3. С. 171–182.
5. Минин В. А., Зацман И. М., Хавансков В. А., Шубников С. К. Определение индикаторов интенсивности переноса научных знаний в сферу технологий // Системы и средства информатики, 2018. Т. 28. № 3. С. 189–203.
6. Зацман И. М., Шубников С. К. Принципы обработки информационных ресурсов для оценки инновационного потенциала направлений научных исследований // Электронные библиотеки: перспективные методы и технологии, электронные коллекции: Труды IX Всероссийской научной конференции RCDL'2007. — Переславль: Университет города Переславля, 2007. С. 35–44.
7. Минин В. А., Зацман И. М., Кружков М. Г., Норекиян Т. П. Методологические основы создания информационных систем для вычисления индикаторов тематических взаимосвязей науки и технологий // Информатика и её применения, 2013. Т. 7. Вып. 1. С. 70–81.
8. Минин В. А., Зацман И. М., Хавансков В. А., Шубников С. К. Архитектурные решения для систем вычисления индикаторов тематических взаимосвязей науки и технологий // Системы и средства информатики, 2013. Т. 23. № 2. С. 260–283.
9. Зацман И. М., Хавансков В. А., Шубников С. К. Метод извлечения библиографической информации из полнотекстовых описаний изобретений // Информатика и её применения, 2013. Т. 7. Вып. 4. С. 52–65.
10. Хавансков В. А., Шубников С. К. Поиск и рубрицирование ссылок на цитируемые публикации в электронных библиотеках полнотекстовых описаний изобретений // Электронные библиотеки: перспективные методы и технологии, электронные коллекции: Труды XVI Всеросс. научной конф. — Дубна: ОИЯИ, 2014. С. 165–173.
11. Минин В. А., Зацман И. М., Хавансков В. А., Шубников С. К. Индикаторы тематических взаимосвязей науки и технологий: от текста к числам // Информатика и её применения, 2014. Т. 8. Вып. 3. С. 114–125.
12. Минин В. А., Зацман И. М., Хавансков В. А., Шубников С. К. Индикаторы тематических взаимосвязей науки и информационно-компьютерных технологий в начале XXI века // Информатика и её применения, 2015. Т. 9. Вып. 2. С. 111–120.
13. Левицкий М. М. Между жидкостью и твердым телом // Химия, 2009. № 11(777). С. 1–11.
14. Brody T. P. Birth of the active matrix // Inform. Display, 1997. Vol. 13. No. 10. P. 28–32.

Поступила в редакцию 15.09.18

THEMATIC INDICATORS OF TRANSFER OF SCIENTIFIC KNOWLEDGE TO THE FIELD OF TECHNOLOGY

V. A. Minin, I. M. Zatsman, V. A. Havanskov, and S. K. Shubnikov

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

Abstract: Indicator assessment of the process of knowledge transfer is an important element in the studies of information linkages between science and technology. Indicators in question characterize the processes of creation of scientific groundwork for development of advanced technologies. In the process of studying information linkages, an experiment has been conducted to calculate the values of indicators of intensity of scientific knowledge transfer to information technology development. To this end, the eleven-year array of full-text descriptions of inventions obtained from Rospatent has been analyzed. The semantic meanings of the intensity indicators that characterize science–technology linkages are considered. As a result of the experiment, the values of indicators that characterize the intensity of the transfer of knowledge into the sphere of information technology development are calculated. The analysis of the obtained results allows drawing a conclusion that the intensity indicators ordered by the frequency of mentions of scientific disciplines make it possible to assess the degree of their influence on the development of information technologies.

Keywords: science–technology linkages; citation of publications in inventions; intensity indicators; indicator assessment; information technologies

DOI: 10.14357/08696527180417

Acknowledgments

The research was supported by the Russian Foundation for Basic Research (project 16-07-00075).

References

1. Minin, V. A., I. M. Zatsman, V. A. Havanskov, and S. K. Shubnikov. 2016. Intensivnost' tsitirovaniya nauchnykh publikatsiy v izobreteniyakh po informatsionno-komp'yuternym tekhnologiyam, patentuemykh v Rossii otechestvennymi i zarubezhnymi zayavitelyami [Intensity of citation of scientific publications in inventions on information and computer technologies patented in Russia by domestic and foreign applicants]. *Informatika i ee Primeneniya — Inform. Appl.* 10(2):107–122.
2. Minin, V. A., I. M. Zatsman, V. A. Khavanskov, and S. K. Shubnikov. 2017. Methods of indicator-based assessment of knowledge transfer from science technology. *Automatic Documentation Math. Linguistics* 51(4):180–190.

3. Zatsman, I. M., G. V. Lukyanov, V. A. Minin, V. A. Havanskov and S. K. Shubnikov. 2017. Indicator evaluation of processes of knowledge transfer from science to technology. *Informatika i ee Primeneniya — Inform. Appl.* 11(3):132–141.
4. Minin, V. A., I. M. Zatsman, V. A. Havanskov, and S. K. Shubnikov. 2017. Make-tirovaniye informatsionnoy sistemy vychisleniya indikatorov perenosa nauchnykh znaniy v sferu tekhnologiy [Prototyping an information system for the calculation of indicators of the intensity of the transfer of scientific knowledge in the field of technology]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 27(3):171–182.
5. Minin, V. A., I. M. Zatsman, V. A. Havanskov, and S. K. Shubnikov. 2018. Opredelenie indikatorov intensivnosti perenosa nauchnykh znaniy v sferu tekhnologiy [Determination of indicators of the intensity of transfer of scientific knowledge in the field of technology]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(3):189–203.
6. Zatsman, I. M., and S. K. Shubnikov. 2007. Printsipy obrabotki informatsionnykh resursov dlya otsenki innovatsionnogo potentsiala napravleniy nauchnykh issledovaniy [The principles of processing of information resources for estimation of innovative potential of the directions of scientific research]. *Tr. 9th Vseross. nauchnoy konf. “Elektronnye biblioteki: perspektivnye metody i tekhnologii, elektronnye kolleksii”* [9th All-Russian Scientific Conference “Digital Libraries: Advanced Methods and Technologies, Digital Collections” Proceedings]. Pereslavl-Zalessky: Pereslavl University. 35–44.
7. Minin, V. A., I. M. Zatsman, M. G. Kruzhek, and T. P. Norekyan. 2013. Metodologicheskie osnovy sozdaniya informatsionnykh sistem dlya vychisleniya indikatorov tematicheskikh vzaimosvyazey nauki i tekhnologiy [Methodological bases for creating information systems calculating indicators of thematic linkages between science and technologies]. *Informatika i ee Primeneniya — Inform. Appl.* 7(1):70–81.
8. Minin, V. A., I. M. Zatsman, V. A. Havanskov, and S. K. Shubnikov. 2013. Arkhitekturnye resheniya dlya sistem vychisleniya indikatorov tematicheskikh vzaimosvyazey nauki i tekhnologiy [Information system conceptual decisions for assessment of linkages between science and technologies]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 23(2):260–283.
9. Zatsman, I. M., V. A. Havanskov, and S. K. Shubnikov. 2013. Metod izvlecheniya bibliograficheskoy informatsii iz polnotekstovykh opisaniy izobreteniy [Method of bibliographic information extraction from full-text descriptions of inventions]. *Informatika i ee Primeneniya — Inform. Appl.* 7(4):52–65.
10. Khavanskov, V. A., and S. K. Shubnikov. 2014. Poisk i rubritirovanie ssylok na tsitiruemye publikatsii v elektronnykh bibliotekakh polnotekstovykh opisaniy izobreteniy [Identification and classification of citation references in digital libraries of full-text patent descriptions]. *Elektronnye biblioteki: perspektivnye metody i tekhnologii, elektronnye kolleksii: Tr. XVI Vseross. nauchnoy konf.* [Digital Libraries: Methods and Technologies, Digital Collections: 16th All-Russian Scientific Conference Proceedings]. Dubna: Joint Institute for Nuclear Research. 165–173.
11. Minin, V. A., I. M. Zatsman, V. A. Havanskov, and S. K. Shubnikov. 2014. Indikatory tematicheskikh vzaimosvyazey nauki i tekhnologiy: ot teksta k chislam [Indicators of thematic relationships between science and technology: From text to numbers]. *Informatika i ee Primeneniya — Inform. Appl.* 8(3):111–120.

12. Minin, V. A., I. M. Zatsman, V. A. Havanskov, and S. K. Shubnikov. 2015. Indikatory tematiceskikh vzaimosvyazey nauki i informatsionno-komp'yuternykh tekhnologiy v nachale XXI veka [Indicators for thematic linkages between science and information and computer technologies at the beginning of the XXI century]. *Informatika i ee Primeneniya — Inform. Appl.* 9(2):111–120.
13. Levitskiy, M. M. 2009. Mezhdru zhidkost'yu i tverdym telom [Between liquid and solid]. *Khimiya* [Chemistry] 11(777):1–11.
14. Brody, T. P. 1997. Birth of the active matrix. *Inform. Display* 13(10):28–32.

Received September 15, 2018

Contributors

Minin Vladimir A. (b. 1941) — Doctor of Science in physics and mathematics, consultant, Institute of Informatics Problems, Federal Research Center “Computer Sciences and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; aleksiss@ya.ru

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

Havanskov Valeriy A. (b. 1950) — scientist, Institute of Informatics Problems, Federal Research Center “Computer Sciences and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; chavanskov@yandex.ru

Shubnikov Sergey K. (b. 1955) — senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Sciences and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; sergeysh50@yandex.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

Архипов Олег Петрович (р. 1948) — директор Орловского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

Башлыкова Анна Александровна (р. 1989) — кандидат технических наук, доцент Московского технологического университета (МИРЭА)

Босов Алексей Вячеславович (р. 1969) — доктор технических наук, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Быстров Игорь Иванович (р. 1931) — доктор технических наук, профессор, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Веселов Виталий Николаевич (р. 1940) — кандидат технических наук, ведущий инженер Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

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

Волович Константин Иосифович (р. 1970) — старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Гасилов Артур Владимирович (р. 1992) — младший научный сотрудник Орловского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

Грушо Александр Александрович (р. 1946) — доктор физико-математических наук, профессор, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Захаров Виктор Николаевич (р. 1948) — доктор технических наук, доцент, ученый секретарь Федерального исследовательского центра «Информатика и управление» Российской академии наук

Зацаринный Александр Алексеевич (р. 1951) — доктор технических наук, заместитель директора Федерального исследовательского центра «Информатика и управление» Российской академии наук (ФИЦ ИА РАН); главный научный сотрудник Института проблем информатики ФИЦ ИУ РАН

Зацман Игорь Моисеевич (р. 1952) — доктор технических наук, заведующий отделом Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Ильин Александр Владимирович (р. 1975) — кандидат технических наук, ведущий инженер Государственного научно-исследовательского института авиационных систем

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

Калашников Сергей Тимофеевич (р. 1956) — кандидат технических наук, заведующий отделом Южно-Уральского научного центра Уральского отделения Российской академии наук

Каменщиков Андрей Александрович (р. 1982) — кандидат технических наук, старший научный сотрудник Института радиотехники и электроники им. В. А. Котельникова Российской академии наук

Кириков Игорь Александрович (р. 1955) — кандидат технических наук, директор Калининградского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

Козлов Сергей Витальевич (р. 1955) — кандидат технических наук, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Колесников Александр Васильевич (р. 1948) — доктор технических наук, профессор Института физико-математических наук и информационных технологий Балтийского федерального университета имени Иммануила Канта; старший научный сотрудник Калининградского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

Колин Константин Константинович (р. 1935) — доктор технических наук, профессор, заслуженный деятель науки РФ, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Кружков Михаил Григорьевич (р. 1975) — старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Кудрявцев Алексей Андреевич (р. 1978) — кандидат физико-математических наук, доцент кафедры математической статистики факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова

Листопад Сергей Викторович (р. 1984) — кандидат технических наук, старший научный сотрудник Калининградского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

Маньяков Юрий Анатольевич (р. 1984) — старший научный сотрудник Орловского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

Минин Владимир Александрович (р. 1941) — доктор физико-математических наук, консультант Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Олейников Александр Яковлевич (р. 1939) — доктор технических наук, главный научный сотрудник Института радиотехники и электроники им. В. А. Котельникова Российской академии наук

Палионная Софья Игоревна (р. 1995) — студентка кафедры математической статистики факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова

Сучков Александр Павлович (р. 1954) — доктор технических наук, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Тимонина Елена Евгеньевна (р. 1952) — доктор технических наук, профессор, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Хавансков Валерий Александрович (р. 1950) — научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Хорошилов Александр Алексеевич (р. 1952) — доктор технических наук, профессор, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Хорошилов Алексей Александрович (р. 1988) — кандидат технических наук, научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Читалов Дмитрий Иванович (р. 1989) — младший научный сотрудник Южно-Уральского научного центра Уральского отделения Российской академии наук

Чусов Игорь Иванович (р. 1945) — кандидат технических наук, старший научный сотрудник Института радиотехники и электроники им. В. А. Котельникова Российской академии наук

Шоргин Всеволод Сергеевич (р. 1978) — кандидат технических наук, старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Шоргин Сергей Яковлевич (р. 1952) — доктор физико-математических наук, профессор, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Шубников Сергей Константинович (р. 1955) — старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Яковлев Олег Альбертович (р. 1992) — младший научный сотрудник Орловского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

АВТОРСКИЙ УКАЗАТЕЛЬ ЗА 2018 г.

	№	Стр.
Абрамов Г. В., Гаврилов А. Н. Математические методы исследования кинетики формирования кластеров углерода в плазме	2	116–127
Адамович И. М., Волков О. И. Линейное упорядочение множества правил в системе извлечения биографических фактов Т-парсер	3	217–226
Адамович И. М., Волков О. И. Модель процесса коррекции ошибок в семантической сети	1	65–76
Адамович И. М., Волков О. И. Элементы самообучения в системе извлечения биографических фактов Т-парсер	4	145–155
Андреева О. В., Багиров М. Б., Данькина А. А., Федорова Т. О., Шевелёва М. М. Интеллектуальный анализ данных на базе Stanford CoreNLP для определения частей речи в русском языке	2	145–153
Архангельская Е. В. Об одной реализации метода динамического программирования для решения задачи о замене оборудования с помощью прикладной программы	2	178–188
Архипов О. П., Гасилов А. В., Маньяков Ю. А., Яковлев О. А. Алгоритм обследования замкнутого помещения автономным мобильным роботом	4	10–21
Багиров М. Б. см. Андреева О. В.		
Башлыкова А. А., Зацаринный А. А., Каменщиков А. А., Козлов С. В., Олейников А. Я., Чусов И. И. Интероперабельность как научно-методическая и нормативная основа бесшовной интеграции информационно-телекоммуникационных систем	4	61–72
Борисов А. В., Босов А. В., Иванов А. В., Корепанов Э. Р. К вопросу о расчете надежности информационно-телекоммуникационных систем: учет характеристик программного обеспечения	1	20–34
Борисов А. В., Босов А. В., Иванов А. В., Стефанович А. И. Анализ функционирования информационно-аналитических систем: методика расчета временных затрат и численности персонала	3	39–53
Борисов А. В., Босов А. В., Иванов А. В., Чавтараев Р. Б. Имитационное моделирование пользовательской активности для оценивания вероятностно-временных характеристик программного обеспечения	2	20–33

	№	Стр.
Босов А. В. см. Борисов А. В.		
Босов А. В. см. Борисов А. В.		
Босов А. В. см. Борисов А. В.		
Босов А. В. см. Сучков А. П.		
Быстров И. И., Веселов В. Н., Колин К. К. Основы живучести автоматизированных организаций информационного общества	4	110–121
Веселов В. Н. см. Быстров И. И.		
Волков О. И. см. Адамович И. М.		
Волков О. И. см. Адамович И. М.		
Волков О. И. см. Адамович И. М.		
Волович К. И. Некоторые системотехнические вопросы предоставления вычислительных ресурсов для научных исследований в гибридной высокопроизводительной облачной среде	4	98–109
Воронин А. В. см. Зацаринный А. А.		
Гаврилов А. Н. см. Абрамов Г. В.		
Гаврилов В. Е., Зацаринный А. А. Некоторые системотехнические вопросы использования интеллектуального анализа данных для обеспечения защиты информации в ситуационных центрах	1	89–98
Гасилов А. В. см. Архипов О. П.		
Гершельман А. Ф. см. Ядринцев В. В.		
Гончаренко М. Б. Оценка формы отклика в сигналах магнитоэнцефалограмм с шумом в виде конечной нормальной смеси	3	72–85
Гончаренко М. Б., Захарова Т. В. Вероятностный подход к решению обратной задачи магнитоэнцефалографии	1	35–52
Гордиенко Д. Е., Косолапов Ю. В., Мышко А. С. Применение нечеткого защищенного хранилища для исправления неточностей в аутентификационных данных	1	156–176
Горшенин А. К., Кузьмин В. Ю. Прогнозирование моментов конечных нормальных смесей с использованием нейронных сетей прямого распространения	3	62–71
Грушо А. А., Забежайло М. И., Грушо Н. А., Тимонина Е. Е. Информационная безопасность на основе метаданных в компонентно-интеграционных архитектурах информационных систем	2	34–41
Грушо А. А., Забежайло М. И., Зацаринный А. А., Николаев А. В., Писковский В. О., Сенчило В. В., Сударилов И. В., Тимонина Е. Е. Об анализе ошибочных состояний в распределенных вычислительных системах	1	99–109

	№	Стр.
Грушо А. А., Тимонина Е. Е., Шоргин С. Я. Возможности преодоления защиты сети, использующей метаданные для управления сетевыми соединениями	4	22–30
Грушо Н. А., Сенчило В. В. Моделирование безопасных архитектур распределенных информационно-вычислительных систем на основе комплексной виртуализации	1	110–122
Грушо Н. А. см. Грушо А. А.		
Гузев О. Ю., Чижов И. В. Балансировка нагрузки в защищенных сетях с использованием технологии SDN	1	123–138
Гузев О. Ю., Чижов И. В. SDN-балансировка нагрузки на криптографические маршрутизаторы при объединении центров обработки данных	1	139–155
Данькина А. А. см. Андреева О. В.		
Дулин С. К., Дулина Н. Г. Использование диссеминационных алгоритмов для формирования неструктурированной текстовой информации в базе геоданных	2	42–59
Дулина Н. Г. см. Дулин С. К.		
Егоров В. Б. Накладные расходы виртуализации и влияющие на них факторы	3	141–152
Забейайло М. И. см. Грушо А. А.		
Забейайло М. И. см. Грушо А. А.		
Замковой А. А. см. Усманова К. Р.		
Захаров В. Н., Хорошилов Ал-др А., Хорошилов Ал-ей А. Автоматическое построение синтаксических моделей языка для систем обработки текстовой информации	4	4–9
Захарова Т. В. см. Гончаренко М. Б.		
Зацаринный А. А., Воронин А. В., Ионенков Ю. С. Особенности оценки эффективности геоинформационной системы как элемента ситуационного центра	2	75–87
Зацаринный А. А., Ионенков Ю. С., Сучков А. П. Некоторые аспекты оценки эффективности облачных технологий	3	104–117
Зацаринный А. А. см. Башлыкова А. А.		
Зацаринный А. А. см. Гаврилов В. Е.		
Зацаринный А. А. см. Грушо А. А.		
Зацман И. М. Методология обратимой генерализации в контексте классификации информационных трансформаций	2	128–144
Зацман И. М. Стадии целенаправленного извлечения знаний, имплицитированных в параллельных текстах	3	175–188
Зацман И. М., Кружков М. Г. Надкорпусная база данных коннекторов: развитие системы терминов проектирования	4	156–167
Зацман И. М. см. Минин В. А.		
Зацман И. М. см. Минин В. А.		

	№	Стр.
Зиганшина Ф. Т. см. Спивак С. И.		
Иванов А. В. см. Борисов А. В.		
Иванов А. В. см. Борисов А. В.		
Иванов А. В. см. Борисов А. В.		
Ильин А. В., Ильин В. Д. Модель распределения потоков в однопродуктовой сетевой системе	2	170–177
Ильин А. В., Ильин В. Д. Ситуационное онлайн-планирование ресурсов по обязательным и ориентирующим правилам	1	177–191
Ильин А. В., Ильин В. Д. Ситуационное планирование производства в сетевой М2М-системе	4	136–144
Ильин В. Д. Технология назначенных платежей в среде цифровых двойников	3	227–235
Ильин В. Д. см. Ильин А. В.		
Ильин В. Д. см. Ильин А. В.		
Ильин В. Д. см. Ильин А. В.		
Инькова О. Ю., Кружков М. Г. Метод описания структуры неоднословных коннекторов в надкорпусных базах данных	4	168–181
Ионенков Ю. С. см. Зацаринный А. А.		
Ионенков Ю. С. см. Зацаринный А. А.		
Исмагилова А. С. см. Спивак С. И.		
Калашников С. Т. см. Читалов Д. И.		
Калёнов Н. Е. Некоторые оценки поисковых возможностей интернет-каталогов российских библиотек	3	236–242
Каменщиков А. А. см. Башлыкова А. А.		
Кириков И. А., Колесников А. В., Листопад С. В. Визуально-образное моделирование эффекта синергии в гибридных интеллектуальных многоагентных системах	3	165–174
Кириков И. А., Колесников А. В., Листопад С. В. Визуальный язык моделирования эффекта синергии в гибридных интеллектуальных многоагентных системах	4	42–53
Клубина К. В. см. Ядринцев В. В.		
Козлов С. В. см. Башлыкова А. А.		
Колесников А. В., Листопад С. В. Модель гибридной интеллектуальной многоагентной системы гетерогенного мышления для информационной подготовки оперативных решений в региональных электрических сетях	4	31–41
Колесников А. В., Листопад С. В., Майгаков Ф. Г. Мета-язык для гибридных интеллектуальных систем визуального управления электрическими сетями: ресурсы, свойства и действия	3	153–164
Колесников А. В. см. Кириков И. А.		
Колесников А. В. см. Кириков И. А.		

	№	Стр.
Колин К. К. см. Быстров И. И.		
Кондрашев В. А. Архитектура системы предоставления сервисов цифровой платформы для научных исследований	3	131–140
Корепанов Э. Р. см. Борисов А. В.		
Корепанов Э. Р. см. Сеницын И. Н.		
Королёв С. П., Романова И. М., Мальковский С. И., Сорочкин А. А. Сервис-ориентированный интерфейс для доступа к научным данным в области исследования и оперативного мониторинга состояния вулканов Камчатки и Северных Курил	2	88–98
Косолапов Ю. В. см. Гордиенко Д. Е.		
Кружков М. Г. см. Зацман И. М.		
Кружков М. Г. см. Инькова О. Ю.		
Кудияров С. П. см. Усманова К. Р.		
Кудрявцев А. А., Палионная С. И., Шоргин В. С. Априорное обратное гамма-распределение в байесовских моделях массового обслуживания	4	54–60
Кудрявцев А. А., Палионная С. И., Шоргин С. Я. Бета-полиномиальные априорные плотности в байесовских моделях надежности	3	54–61
Кузьмин В. Ю. см. Горшенин А. К.		
Леонтьев Д. В. см. Харитонов Д. И.		
Листопад С. В. см. Кириков И. А.		
Листопад С. В. см. Кириков И. А.		
Листопад С. В. см. Колесников А. В.		
Листопад С. В. см. Колесников А. В.		
Майтаков Ф. Г. см. Колесников А. В.		
Мальковский С. И. см. Королёв С. П.		
Маньяков Ю. А. см. Архипов О. П.		
Маргышкин Р. В. см. Усманова К. Р.		
Минин В. А., Зацман И. М., Хавансков В. А., Шубников С. К. Определение индикаторов интенсивности переноса научных знаний в сферу технологий	3	189–203
Минин В. А., Зацман И. М., Хавансков В. А., Шубников С. К. Тематические индикаторы переноса научных знаний в сферу технологий	4	182–198
Мышко А. С. см. Гордиенко Д. Е.		
Низамутдинов М. М., Орешников В. В. Информационная поддержка формирования стратегий инновационного развития регионов на основе адаптивной имитационной модели	2	154–169
Никишин Д. А. Процессы генерализации в аналоговой и цифровой картографии	3	204–216

	№	Стр.
Никишин Д. А. Сопоставление особенностей представления гео- данных в картографии и в геоинформатике	2	60–74
Николаев А. В. см. Грушо А. А.		
Олейников А. Я. см. Башлыкова А. А.		
Орешников В. В. см. Низамутдинов М. М.		
Палионная С. И. см. Кудрявцев А. А.		
Палионная С. И. см. Кудрявцев А. А.		
Писковский В. О. см. Грушо А. А.		
Романова И. М. см. Королёв С. П.		
Сенчило В. В. см. Грушо А. А.		
Сенчило В. В. см. Грушо Н. А.		
Синицын В. И. см. Синицын И. Н.		
Синицын В. И. см. Синицын И. Н.		
Синицын В. И. см. Синицын И. Н.		
Синицын И. Н., Синицын В. И. Аналитическое моделирова- ние нормальных процессов в вольтерровских стохастических системах	2	4–19
Синицын И. Н., Синицын В. И. Аналитическое моделирова- ние распределений с инвариантной мерой в вольтерровских стохастических системах	3	4–25
Синицын И. Н., Синицын В. И., Корепанов Э. Р. Аналитиче- ский синтез субоптимальных фильтров методами моментов	1	4–19
Сорокин А. А. см. Королёв С. П.		
Спивак С. И., Зиганшина Ф. Т., Исмагилова А. С. Инфор- мативность кинетического эксперимента и области неопреде- ленности параметров кинетических моделей	1	77–88
Стефанович А. И. см. Борисов А. В.		
Стрижов В. В. см. Усманова К. Р.		
Судариков И. В. см. Грушо А. А.		
Сучков А. П. Анализ процессов межведомственного информа- ционного взаимодействия	3	118–130
Сучков А. П. Процессная модель модернизации и развития информационных систем на всех стадиях жизненного цикла	4	86–97
Сучков А. П., Босов А. В. Концептуальные подходы к созда- нию информационно-аналитического ситуационного центра в интересах стратегического планирования в области обеспе- чения национальной безопасности	4	73–85
Сучков А. П. см. Зацаринный А. А.		
Тарасов Г. В. см. Харитонов Д. И.		
Тимонина Е. Е. см. Грушо А. А.		
Тимонина Е. Е. см. Грушо А. А.		
Тимонина Е. Е. см. Грушо А. А.		

	№	Стр.
Тихомиров И. А. см. Ядринцев В. В.		
Усманова К. Р., Кудряков С. П., Мартышкин Р. В., Замковой А. А., Стрижов В. В. Анализ зависимостей между показателями при прогнозировании объема грузоперевозок	3	86–103
Федорова Т. О. см. Андреева О. В.		
Хавансков В. А. см. Минин В. А.		
Хавансков В. А. см. Минин В. А.		
Харитонов Д. И., Тарасов Г. В., Леонтьев Д. В. Программная реализация метода генерации сетей Петри большого размера	2	99–115
Хорошилов Ал-др А. см. Захаров В. Н.		
Хорошилов Ал-ей А. см. Захаров В. Н.		
Чавтараев Р. Б. см. Борисов А. В.		
Чижов И. В. см. Гузев О. Ю.		
Чижов И. В. см. Гузев О. Ю.		
Читалов Д. И., Калашников С. Т. Разработка приложения для подготовки расчетных сеток с градуирующими и изогнутыми краями для программной среды OpenFOAM	4	122–135
Чусов И. И. см. Башлыкова А. А.		
Шевелёва М. М. см. Андреева О. В.		
Шоргин В. С. Кудрявцев А. А.		
Шоргин С. Я. см. Грушо А. А.		
Шоргин С. Я. см. Кудрявцев А. А.		
Шубников С. К. см. Минин В. А.		
Шубников С. К. см. Минин В. А.		
Ядринцев В. В., Клубина К. В., Тихомиров И. А., Гершельман А. Ф. Выбор серверного решения для цифровой платформы поиска и анализа текстов	3	26–38
Яковлев О. А. Методы инициализации воксельного объема в задаче трехмерной реконструкции	1	53–64
Яковлев О. А. см. Архипов О. П.		

2018 AUTHOR INDEX

	No.	Page
Abramov G. V. and Gavrilov A. N. Mathematical methods for studying the kinetics of formation of carbon clusters in plasma	2	116–127
Adamovich I. M. and Volkov O. I. Elements of machine learning in the T-parser system of facts extraction	4	145–155
Adamovich I. M. and Volkov O. I. Linear ordering of the rules set in the T-parser system of facts extraction	3	217–226
Adamovich I. M. and Volkov O. I. The model of semantic net error correction process	1	65–76
Andreeva O. V., Bagirov M. B., Dankina A. A., Fedorova T. O., and Sheveleva M. M. Intellectual analysis of data on the basis of Stanford CoreNLP for POS tagging of texts in the Russian language	2	145–153
Arkhangelskaya E. V. About an implementation of the method of dynamic programming to solve the problem of equipment replacement with an application program	2	178–188
Arkhipov O. P., Gasilov A. V., Maniakov Yu. A., and Yakovlev O. A. Enclosed room exploration algorithm for an autonomous mobile robot	4	10–21
Bagirov M. B. see Andreeva O. V.		
Bashlykova A. A., Zatsarinny A. A., Kamenshchikov A. A., Kozlov S. V., Oleynikov A. Ya., and Chusov I. I. Interoperability as a scientific, methodical, and regulatory base of information and telecommunication systems seamless integration	4	61–72
Borisov A. V., Bosov A. V., Ivanov A. V., and Chavtarav R. B. Monte Carlo based user activity simulation for software performance evaluation	2	20–33
Borisov A. V., Bosov A. V., Ivanov A. V., and Korepanov E. R. To the reliability of an information-telecommunication system: An approach to recognition of reliable software characteristics	1	20–34
Borisov A. V., Bosov A. V., Ivanov A. V., and Stefanovich A. I. Analytical information systems performance analysis: Methodology for timetable and staff quantity evaluation	3	39–53
Bosov A. V. see Borisov A. V.		
Bosov A. V. see Borisov A. V.		

	No.	Page
Bosov A. V. see Borisov A. V.		
Bosov A. V. see Suchkov A. P.		
Bystrov I. I., Veselov V. N., and Kolin K. K. Information society automated organizations survivability basics	4	110–121
Chavtaraev R. B. see Borisov A. V.		
Chitalov D. I. and Kalashnikov S. T. Development of an application for preparation of computational meshes with graduating and curved edges for the OpenFOAM software	4	122–135
Chizhov I. V. see Guzev O. Yu.		
Chizhov I. V. see Guzev O. Yu.		
Chusov I. I. see Bashlykova A. A.		
Dankina A. A. see Andreeva O. V.		
Dulin S. K. and Dulina N. G. Use of dissemination algorithms for formation of nonstructured text information in the geodata base	2	42–59
Dulina N. G. see Dulin S. K.		
Egorov V. B. Virtualization overheads and factors affecting them	3	141–152
Fedorova T. O. see Andreeva O. V.		
Gasilov A. V. see Arkhipov O. P.		
Gavrilov A. N. see Abramov G. V.		
Gavrilov V. E. and Zatsarinny A. A. Regarding systemic and technical problems of applying intellectual data analysis for providing information protection in situational centers	1	89–98
Gershel'man A. F. see Yadrintsev V. V.		
Goncharenko M. B. Evaluation of the waveform in the magnetoencephalographic signals with noise in the form of a finite normal mixture	3	72–85
Goncharenko M. B. and Zakharova T. V. Probabilistic approach to solving the magnetoencephalography inverse problem	1	35–52
Gordienko D. E., Kosolapov Yu. V., and Mishko A. S. Using the fuzzy vault to correct inaccuracies in authentication data	1	156–176
Gorshenin A. K. and Kuzmin V. Yu. Forecasting moments of finite normal mixtures using feedforward neural networks	3	62–71
Grusho A. A., Timonina E. E., and Shorgin S. Ya. Overcoming of protection of a network where connections are controlled by meta data	4	22–30
Grusho A. A., Zabezhailo M. I., Grusho N. A., and Timonina E. E. Information security on the basis of meta data in enterprise application integration architecture of information systems	2	34–41

	No.	Page
Grusho A. A., Zabezhailo M. I., Zatsarinny A. A., Nikolaev A. V., Piskovski V. O., Senchilo V. V., Sudarikov I. V., and Timonina E. E. About the analysis of erratic statuses in the distributed computing systems	1	99–109
Grusho N. A. and Senchilo V. V. Modeling of secure architecture of distributed information systems on the basis of integrated virtualization	1	110–122
Grusho N. A. see Grusho A. A.		
Guzev O. Yu. and Chizhov I. V. SDN load balancing for secure networks	1	123–138
Guzev O. Yu. and Chizhov I. V. SDN load balancing on L3-VPN gateways in data centers interconnection	1	139–155
Havanskov V. A. see Minin V. A.		
Havanskov V. A. see Minin V. A.		
Ilyin A. V. and Ilyin V. D. Situational online resource planning in accordance with mandatory and orienting rules	1	177–191
Ilyin A. V. and Ilyin V. D. Situational production planning in a network M2M-system	4	136–144
Ilyin A. V. and Ilyin V. D. The model of flows distribution in a single-product network system	2	170–177
Ilyin V. D. Designated payments technology in digital twins environment	3	227–235
Ilyin V. D. see Ilyin A. V.		
Ilyin V. D. see Ilyin A. V.		
Ilyin V. D. see Ilyin A. V.		
Inkova O. Yu. and Kruzhkov M. G. Method for description of multiword connectives in Supracorpora databases	4	168–181
Ionenkov Yu. S. see Zatsarinny A. A.		
Ionenkov Yu. S. see Zatsarinny A. A.		
Ismagilova A. S. see Spivak S. I.		
Ivanov A. V. see Borisov A. V.		
Ivanov A. V. see Borisov A. V.		
Ivanov A. V. see Borisov A. V.		
Kalashnikov S. T. see Chitalov D. I.		
Kalenov N. E. Some evaluation of the Russian libraries online catalogs search capabilities	3	236–242
Kamenshchikov A. A. see Bashlykova A. A.		
Kharitonov D. I., Tarasov G. V., and Leontyev D. V. Software implementation of a method for generation of Petri nets of large size	2	99–115
Khoroshilov Al-dr A. see Zakharov V. N.		
Khoroshilov Al-ey A. see Zakharov V. N.		

	No.	Page
Kirikov I. A., Kolesnikov A. V., and Listopad S. V. Visual language for modeling of synergy effect in hybrid intelligent multiagent systems	4	42–53
Kirikov I. A., Kolesnikov A. V., and Listopad S. V. Visual simulation of synergy effect in hybrid intelligent multiagent systems	3	165–174
Klyubina K. V. see Yadrintsev V. V.		
Kolesnikov A. V. and Listopad S. V. Model of a hybrid intelligent multiagent system of heterogeneous thinking for preparation of information about operational decisions in a regional power system	4	31–41
Kolesnikov A. V., Listopad S. V., and Maitakov F. G. Met-alanguage for hybrid intelligent systems for visual control of electrical networks: Resources, properties, and actions	3	153–164
Kolesnikov A. V. see Kirikov I. A.		
Kolesnikov A. V. see Kirikov I. A.		
Kolin K. K. see Bystrov I. I.		
Kondrashev V. A. Architecture of the service delivery system for the research services digital platform	3	131–140
Korepanov E. R. see Borisov A. V.		
Korepanov E. R. see Sinitsyn I. N.		
Korolev S. P., Romanova I. M., Malkovsky S. I., and Sorokin A. A. Service-oriented interface to access scientific data for study and state operational monitoring of volcanoes of Kamchatka and Northern Kuriles	2	88–98
Kosolapov Yu. V. see Gordienko D. E.		
Kozlov S. V. see Bashlykova A. A.		
Kruzhkov M. G. see Inkova O. Yu.		
Kruzhkov M. G. see Zatsman I. M.		
Kudiyarov S. P. see Usmanova K. R.		
Kudryavtsev A. A., Palionnaia S. I., and Shorgin S. Ya. Beta-polynomial <i>a priori</i> densities in Bayesian reliability models	3	54–61
Kudryavtsev A. A., Palionnaia S. I., and Shorgin V. S. <i>A priori</i> inverse gamma distribution in Bayesian queuing models	4	54–60
Kuzmin V. Yu. see Gorshenin A. K.		
Leontyev D. V. see Kharitonov D. I.		
Listopad S. V. see Kirikov I. A.		
Listopad S. V. see Kirikov I. A.		
Listopad S. V. see Kolesnikov A. V.		
Listopad S. V. see Kolesnikov A. V.		
Maitakov F. G. see Kolesnikov A. V.		

	No.	Page
Malkovsky S. I. see Korolev S. P.		
Maniakov Yu. A. see Arkhipov O. P.		
Martyshkin R. V. see Usmanova K. R.		
Minin V. A., Zatsman I. M., Havanskov V. A., and Shubnikov S. K. Calculation of intensity indicators of science-technology linkages	3	189–203
Minin V. A., Zatsman I. M., Havanskov V. A., and Shubnikov S. K. Thematic indicators of transfer of scientific knowledge to the field of technology	4	182–198
Mishko A. S. see Gordienko D. E.		
Nikishin D. A. Comparison of the features of the representation of geodata in cartography and geoinformatics	2	60–74
Nikishin D. A. Processes of generalization in analog and digital cartography	3	204–216
Nikolaev A. V. see Grusho A. A.		
Nizamutdinov M. M. and Oreshnikov V. V. Conception of a decision support system for regions innovative development management based on the adaptive simulation model	2	154–169
Oleynikov A. Ya. see Bashlykova A. A.		
Oreshnikov V. V. see Nizamutdinov M. M.		
Palionnaia S. I. see Kudryavtsev A. A.		
Palionnaia S. I. see Kudryavtsev A. A.		
Piskovski V. O. see Grusho A. A.		
Romanova I. M. see Korolev S. P.		
Senchilo V. V. see Grusho A. A.		
Senchilo V. V. see Grusho N. A.		
Sheveleva M. M. see Andreeva O. V.		
Shorgin S. Ya. see Grusho A. A.		
Shorgin S. Ya. see Kudryavtsev A. A.		
Shorgin V. S. see Kudryavtsev A. A.		
Shubnikov S. K. see Minin V. A.		
Shubnikov S. K. see Minin V. A.		
Sinitsyn I. N. and Sinitsyn V. I. Analytical modeling of distributions with invariant measure in Volterra stochastic systems	3	4–25
Sinitsyn I. N. and Sinitsyn V. I. Analytical modeling of normal processes in Volterra stochastic systems	2	4–19
Sinitsyn I. N., Sinitsyn V. I., and Korepanov E. R. Analytical synthesis of suboptimal filters by moments methods	1	4–19
Sinitsyn V. I. see Sinitsyn I. N.		
Sinitsyn V. I. see Sinitsyn I. N.		
Sinitsyn V. I. see Sinitsyn I. N.		
Sorokin A. A. see Korolev S. P.		

	No.	Page
Spivak S. I., Ziganshina F. T., and Ismagilova A. S. Informativity of a kinetic experiment and uncertainty regions of kinetic model	1	77–88
Stefanovich A. I. see Borisov A. V.		
Strijov V. V. see Usmanova K. R.		
Suchkov A. P. Analysis of the interdepartmental information exchange processes	3	118–130
Suchkov A. P. Process model of information systems modernization and development at all stages of their life cycle	4	86–97
Suchkov A. P. and Bosov A. V. Conceptual approaches to creating an information and analytical situational center for strategic planning in the field of national security	4	73–85
Suchkov A. P. see Zatsarinny A. A.		
Sudarikov I. V. see Grusho A. A.		
Tarasov G. V. see Kharitonov D. I.		
Tikhomirov I. A. see Yadrintsev V. V.		
Timonina E. E. see Grusho A. A.		
Timonina E. E. see Grusho A. A.		
Timonina E. E. see Grusho A. A.		
Usmanova K. R., Kudiyarov S. P., Martyshkin R. V., Zamkovoy A. A., and Strijov V. V. Analysis of relationships between indicators in forecasting cargo transportation	3	86–103
Veselov V. N. see Bystrov I. I.		
Volkov O. I. see Adamovich I. M.		
Volkov O. I. see Adamovich I. M.		
Volkov O. I. see Adamovich I. M.		
Volovich K. I. Organization of calculations in a hybrid high-performance computing cluster for parallel execution of heterogeneous tasks	4	98–109
Voronin A. V. see Zatsarinny A. A.		
Yadrintsev V. V., Klyubina K. V., Tikhomirov I. A., and Gershel'man A. F. Choosing server solution for digital platform of text search and analysis	3	26–38
Yakovlev O. A. Initial bounding box estimation methods for volumetric three-dimensional reconstruction	1	53–64
Yakovlev O. A. see Arkhipov O. P.		
Zabekhailo M. I. see Grusho A. A.		
Zabekhailo M. I. see Grusho A. A.		
Zakharov V. N., Khoroshilov Al-dr A., and Khoroshilov Al-ey A. Automatic construction of syntactic language models for text processing systems	4	4–9
Zakharova T. V. see Goncharenko M. B.		

	No.	Page
Zamkovoy A. A. see Usmanova K. R.		
Zatsarinny A. A., Ionenkov Yu. S., and Suchkov A. P. Some aspects of cloud computing efficiency estimation	3	104–117
Zatsarinny A. A., Voronin A. V., and Ionenkov Yu. S. The features of efficiency evaluation of a geoinformation system as an element of a situational center	2	75–87
Zatsarinny A. A. see Bashlykova A. A.		
Zatsarinny A. A. see Gavrilov V. E.		
Zatsarinny A. A. see Grusho A. A.		
Zatsman I. M. Methodology of reversible generalization in context of classification of information transformations	2	128–144
Zatsman I. M. Stages of goal-oriented discovery of knowledge implied in parallel texts	3	175–188
Zatsman I. M. and Kruzhkov M. G. Supracorpora database of connectives: Design-oriented evolution of the term system	4	156–167
Zatsman I. M. see Minin V. A.		
Zatsman I. M. see Minin V. A.		
Ziganshina F. T. see Spivak S. I.		

Правила подготовки рукописей статей для публикации в журнале «Системы и средства информатики»

Журнал «Системы и средства информатики» публикует теоретические, обзорные и дискуссионные статьи, посвященные научным исследованиям и разработкам в области информационных технологий.

Журнал издается на русском языке. По специальному решению редколлегии отдельные статьи могут печататься на английском языке.

Тематика журнала охватывает следующие направления:

- информационно-телекоммуникационные системы и средства их построения;
- архитектура и программное обеспечение вычислительных машин, комплексов и сетей;
- методы и средства защиты информации.

1. В журнале печатаются статьи, содержащие результаты, ранее не опубликованные и не предназначенные к одновременной публикации в других изданиях.

Публикация предоставленной автором(ами) рукописи не должна нарушать положений глав 69, 70 раздела VII части IV Гражданского кодекса, которые определяют права на результаты интеллектуальной деятельности и средства индивидуализации, в том числе авторские права, в РФ.

Ответственность за нарушение авторских прав, в случае предъявления претензий к редакции журнала, несут авторы статей.

Направляя рукопись в редакцию, авторы сохраняют свои права на данную рукопись и при этом передают учредителям и редколлегии журнала неисключительные права на издание статьи на русском языке (или на языке статьи, если он отличен от русского) и на перевод ее на английский язык, а также на ее распространение в России и за рубежом. Каждый автор должен представить в редакцию подписанный с его стороны «Лицензионный договор о передаче неисключительных прав на использование произведения», текст которого размещен по адресу <http://www.ipiran.ru/publications/licence.doc>. Этот договор может быть представлен в бумажном (в 2-х экз.) или в электронном виде (отсканированная копия заполненного и подписанного документа).

Редколлегия вправе запросить у авторов экспертное заключение о возможности публикации представленной статьи в открытой печати.

2. К статье прилагаются данные автора (авторов) (см. п. 8). При наличии нескольких авторов указывается фамилия автора, ответственного за переписку с редакцией.
3. Редакция журнала осуществляет экспертизу присланных статей в соответствии с принятой в журнале процедурой рецензирования.

Возвращение рукописи на доработку не означает ее принятия к печати.

Доработанный вариант с ответом на замечания рецензента необходимо прислать в редакцию.

4. Решение редколлегии о публикации статьи или ее отклонении сообщается авторам. Редколлегия может также направить авторам текст рецензии на их статью. Дискуссия по поводу отклоненных статей не ведется.

5. Редактура статей высылается авторам для просмотра. Замечания к редакции должны быть присланы авторами в кратчайшие сроки.
6. Рукопись предоставляется в электронном виде в форматах MS WORD (.doc или .docx) или L^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 15 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/2012_22_02_rus/authors.asp и
http://www.ipiran.ru/journal/collected/2012_22_02_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. *Rus. 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 Publs. 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. — N.Y.: Wiley, 1974. 521 p.*)

2. В англоязычной части:

Timoshenko, S. P., D. H. Young, and W. Weaver. 1974. *Vibration problems in engineering*. 4th ed. N.Y.: 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 modelirovaniye plazmy v sisteme kompaktnyy tor [Mathematical modeling of the plasma in the compact torus]. D.Sc. Diss. Moscow. 272 p.

Kozhunova, O. S. 2009. Tekhnologiya razrabotki semanticheskogo slovarya informatsionnogo 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: rust@ipiran.ru (Сейфуль-Мулюков Рустем Бадриевич)
<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 15 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/2012.22.02_rus/authors.asp and
http://www.ipiran.ru/journal/collected/2012.22.02_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. *Rus. 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 slovary informat-sionnogo monitoringa [Technology of development of semantic dictionary of information monitoring system]. PhD Thesis. Moscow: IPI RAN. 23 p.

State standards and patents:

GOST 8.586.5-2005. 2007. Metodika vypolneniya izmereniy. Izmerenie raskhoda i kolichstva zhidkostey i gazov s pomoshch'yu standartnykh suzhayushchikh ustroystv [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. M.: Standardinform Publs. 10 p.

Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. Sposob orientirovaniya po krenu letatel'nogo apparata s opticheskoy golovkoy samonavedeniya [The way to orient on the roll of aircraft with optical homing head]. Patent RF No. 2280590.

References in Latin transcription are presented in the original language.

References in the text are numbered according to the order of their first appearance; the number is placed in square brackets. All items from the reference list should be cited.

10. Manuscripts and additional materials are not returned to Authors by the Editorial Board.
11. Submissions of files by e-mail must include:
 - the journal title and author's name in the "Subject" field;
 - an article and additional materials have to be attached using the "attach" function;
 - an electronic version of the article should contain the file with the text and a separate file with figures.
12. "System and Means of Informatics" journal is not a profit publication. There are no charges for the authors as well as there are no royalties.

Editorial Board address:

FRC CSC RAS, 44, block 2, Vavilov Str., Moscow 119333, Russia

Ph.: +7 (499) 135 86 92, Fax: +7 (495) 930 45 05

e-mail: rust@ipiran.ru (to Prof. Rustem Seyful-Mulyukov)

http://www.ipiran.ru/english/journal_systems.asp

SYSTEMS AND MEANS OF INFORMATICS (SISTEMY I SREDSTVA INFORMATIKI)

SCIENTIFIC JOURNAL

Volume 28 No.4 Year 2018

Editor-in-Chief and Chair of Editorial Council
Academician I. A. Sokolov

IN THIS ISSUE:

- AUTOMATIC CONSTRUCTION OF SYNTACTIC LANGUAGE MODELS
FOR TEXT PROCESSING SYSTEMS
V. N. Zakharov, Al-dr A. Khoroshilov, and Al-ey A. Khoroshilov 4
- ENCLOSED ROOM EXPLORATION ALGORITHM FOR AN AUTONOMOUS MOBILE ROBOT
O. Arkhipov, A. Gasilov, Yu. Maniakov, and O. Yakovlev 10
- OVERCOMING OF PROTECTION OF A NETWORK WHERE CONNECTIONS
ARE CONTROLLED BY META DATA
A. A. Grusho, E. E. Timonina, and S. Ya. Shorgin 22
- MODEL OF A HYBRID INTELLIGENT MULTIAGENT SYSTEM
OF HETEROGENEOUS THINKING FOR PREPARATION OF INFORMATION
ABOUT OPERATIONAL DECISIONS IN A REGIONAL POWER SYSTEM
A. V. Kolesnikov and S. V. Listopad 31
- VISUAL LANGUAGE FOR MODELING OF SYNERGY EFFECT
IN HYBRID INTELLIGENT MULTIAGENT SYSTEMS
I. A. Kirikov, A. V. Kolesnikov and S. V. Listopad 42
- A *PRIORI* INVERSE GAMMA DISTRIBUTION IN BAYESIAN QUEUING MODELS
A. A. Kudryavtsev, S. I. Palionnaia, and V. S. Shorgin 54
- INTEROPERABILITY AS A SCIENTIFIC, METHODICAL, AND REGULATORY BASE
OF INFORMATION AND TELECOMMUNICATION SYSTEMS SEAMLESS INTEGRATION
*A. A. Bashlykova, A. A. Zatsarinny, A. A. Kamenshchikov, S. V. Kozlov,
A. Ya. Oleynikov, and I. I. Chusov* 61
- CONCEPTUAL APPROACHES TO CREATING AN INFORMATION
AND ANALYTICAL SITUATIONAL CENTER FOR STRATEGIC PLANNING
IN THE FIELD OF NATIONAL SECURITY
A. P. Suchkov and A. V. Bosov 73
- PROCESS MODEL OF INFORMATION SYSTEMS MODERNIZATION
AND DEVELOPMENT AT ALL STAGES OF THEIR LIFE CYCLE
A. P. Suchkov 86