

СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ

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

Издается с 1989 года
Журнал выходит ежеквартально

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

РЕДАКЦИОННЫЙ СОВЕТ

академик РАН И. А. Соколов — председатель Редакционного совета
академик РАН Г. И. Савин
академик РАН А. Л. Стемпковский
член-корреспондент РАН Ю. Б. Зубарев
профессор Ш. Долев (S. Dolev, Beer-Sheva, Israel)
профессор Ю. Кабанов (Yu. Kabanov, Besancon, France)
профессор В. Ротарь (V. Rotar, San-Diego, USA)
профессор М. Финкельштейн (M. Finkelstein, Bloemfontein, South Africa)

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

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

Редакция

к.ф.-м.н. Е. Н. Арутюнов
к.ф.-м.н. Р. В. Разумчик
С. Н. Стригина

© Федеральный исследовательский центр «Информатика
и управление» Российской академии наук, 2021

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 31 № 2 Год 2021

СОДЕРЖАНИЕ

«Закладки» без вредоносного кода А. А. Грушо, Н. А. Грушо, М. И. Забежайло, Е. Е. Тимонина	4
Аналитическая модель протокола множественного доступа с прослушиванием канала для приложений индустриального интернета вещей А. С. Царев, Э. М. Хайров, Ю. В. Гайдамака, С. Я. Шоргин	16
Обеспечение надежности центра коллективного пользования ФИЦ ИУ РАН А. А. Зацаринный, А. И. Гаранин, С. А. Денисов	26
Методы анализа данных электроэнцефалографии с применением сверточных и рекуррентных нейронных сетей И. А. Шанин, С. А. Ступников	36
Стимуляция конфликтов агентов в гибридных интеллектуальных многоагентных системах С. В. Листопад, И. А. Кириков	47
Стратегия обследования помещений для автономного мобильного робота на основе карты ракурсов О. А. Яковлев, П. О. Архипов	59
Некоторые вопросы программного определения хранилища данных В. Б. Егоров	70
Информационные аспекты обеспечения безопасности на транспорте: поиск и отбор информации А. В. Борисов, А. В. Босов, Д. В. Жуков, А. В. Иванов	80
Стратегия исследований и разработок в области искусственного интеллекта II: сравнительный анализ наукометрических показателей в мире и в Российской Федерации А. В. Борисов, А. В. Босов, Д. В. Жуков	89

СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ

Том 31 № 2 Год 2021

СОДЕРЖАНИЕ

Аналитические свойства и аспекты вычисления гамма-экспоненциальной функции М. О. Воронцов, А. А. Кудрявцев, С. Я. Шоргин	108
Вариант концептуальной схемы базы геоданных с поддержкой обратимой генерализационной связанности моделей геообъектов Д. А. Никишин	119
Наследуемые лексикографические ресурсы базы данных фразеологического словаря В. В. Вакуленко, И. М. Зацман	129
Индикаторная оценка нестабильности нейронного машинного перевода А. Ю. Егорова, И. М. Зацман, М. Г. Кружков, В. А. Нуриев	139
Устойчивость технологии поддержки конкретно-исторических исследований к попыткам искажения истории И. М. Адамович, О. И. Волков	152
Ситуационная цифровизация товарно-денежного обращения А. В. Ильин, В. Д. Ильин	163
Математическое и компьютерное моделирование агроценоза хлопчатника с учетом возрастной структуры и с произвольными трофическими функциями Р. Н. Одинаев, А. Б. Гафоров	173
Поправка к статье О. М. Корчажиной «SIR-модель как инструмент исследования деструктивных процессов при усвоении нового знания» (Системы и средства информатики, 2021. Т. 31. № 1. С. 168–180)	184
Об авторах	185
Правила подготовки рукописей статей	189
Requirements for manuscripts	193

«ЗАКЛАДКИ» БЕЗ ВРЕДНОСНОГО КОДА*

А. А. Грушо¹, Н. А. Грушо², М. И. Забейло³, Е. Е. Тимонина⁴

Аннотация: Отражен ряд проблем, связанных с угрозами распределенным информационным системам (ИС) из-за наличия BackDoor. Использование BackDoor для атаки на распределенные компьютерные системы обладает большими преимуществами по сравнению с традиционным использованием вредоносного кода. В самом деле, использование BackDoor трудно идентифицировать как специальную атаку, т. е. отличить от ошибки в вычислительных процессах. Однако построение атак с помощью BackDoor значительно сложнее, чем построение атак с помощью вредоносного кода. Построен пример BackDoor с использованием исходных данных информационной технологии (ИТ). Такой подход к построению BackDoor не встречался ни в одном из описаний BackDoor. В работе показано, что существующие технологии поиска уязвимостей в программно-аппаратной системе не решают проблемы поиска BackDoor. Математические модели и методы решения задач поиска BackDoor, кроме частных случаев, еще не разработаны. Принципиально, что BackDoor не связаны с внедрением в компьютерную систему вредоносного кода и с его использованием для нанесения ущерба. Поэтому, как правило, механизм нанесения ущерба через BackDoor отличается от действия вредоносного кода.

Ключевые слова: информационная безопасность; BackDoor; BackDoor через данные

DOI: 10.14357/08696527210201

1 Введение

Всем специалистам по информационной безопасности (ИБ) известно понятие BackDoor (скрытая дверь, далее — СД). Опасность внедрения вредоносного кода связана с нарушением конфиденциальности, целостности и доступности. BackDoor «не крадет» конфиденциальную информацию, не нарушает целостность информации, не определяет доступность информации. Однако каждая

* Работа частично поддержана РФФИ (проект 18-29-03081).

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

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

³ Вычислительный центр им. А. А. Дородницына Федерального исследовательского центра «Информатика и управление» Российской академии наук, m.zabzhailo@yandex.ru

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

их перечисленных угроз может быть реализована только через предоставление доступа к некоторым параметрам ИС.

Работа посвящена описанию некоторых примеров СД, которые отражают основные проблемы угроз, порождаемых существованием СД.

2 Особенности угроз BackDoor

Чаще всего под СД понимается возможность скрытого проникновения в компьютерную систему и внедрение вредоносного кода [1, 2]. В данной работе под СД будем понимать шлюз для воздействия извне на значение конкретного параметра компьютерной системы, т. е. на изменение значений этого параметра при функционировании компьютерной системы. При этом будем допускать, что СД может работать косвенным образом, а именно: в ходе работы компьютерной системы будут созданы условия, позволяющие «сбросить» значение целевого параметра в значение, не соответствующее требуемой конфигурации.

Оказывается, что воздействия через СД на некоторые параметры могут не просто порождать ошибки и сбои, но и существенно влиять на работу больших (критических) систем.

Приведем примеры таких параметров и их влияние, не акцентируя внимание на СД.

1. Экспериментальное облако на платформе OpenStack было полностью выведено из строя за счет изменения (по неизвестной причине) значения одного конфигурационного параметра [3, 4]. При этом поиск ошибки был затруднен, так как все процессы были индексированы как ошибочные.
2. Экспериментально-расчетная система остановилась, когда счетчик числа циклов был сброшен (по неизвестной причине) на одну единицу в минус, т. е. были выполнены положенные n циклов расчетов, но счетчик показывал, что выполнены $n - 1$ циклов. Система остановилась в ожидании окончания n -го цикла.

Использование СД для атаки на распределенные компьютерные системы обладает большими преимуществами по сравнению с традиционным использованием вредоносного кода. В самом деле, использование СД трудно идентифицировать как специальную атаку, т. е. отличить от ошибки в вычислительных процессах. Однако построение атак с помощью СД значительно сложнее, чем построение атак с помощью вредоносного кода.

При работе с СД возникают следующие основные проблемы.

1. Нахождение путей порождения сбоев и ошибок с помощью воздействия на параметр извне компьютерной системы.
2. Наибольший ущерб от атаки может быть достигнут с помощью распространения ошибки, инициированной через СД.

Поиск СД предполагает детальное знание компьютерной системы и относится к трудным переборным задачам. Задача, связанная с распространением ошибки, исследовалась при создании критических информационных ИТ (автоматизированные системы управления технологическими процессами и пр.). Однако, как показывает знакомство с результатами исследований в этой области, возникающие здесь задачи также не из простых.

3 Примеры BackDoors

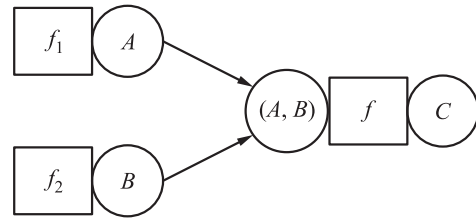
Пример 1. В работе [5] приведены результаты исследований и доказано существование СД в тестовых областях микропроцессоров 20 ведущих производителей. BackDoor использует особенности стандарта JTAG, позволяющие получить доступ ко всем разделам памяти микропроцессора.

Пример 2. Системные администраторы распределенных ИС часто сталкиваются с проблемами, порожденными возможностью запуска процессов, забирающих почти все ресурсы оперативной памяти и центрального процессора. При этом образуются очереди, которые блокируют значительные области таких распределенных систем.

Пример 3. Наиболее частые конфликты с конфигурацией программного обеспечения (ПО) возникают при его обновлении. Поэтому обновления ПО часто порождают СД. Рассмотрим следующий пример. В организации установлен сервер с операционной системой Windows Server 2003. Сервер используется в качестве контроллера домена. В процессе планового усовершенствования было решено перейти на операционную систему Windows Server 2008 R2. Перенос контроллера домена был проведен, однако возникли ошибки в конфигурации контроллера домена, из-за чего пользователи не смогли зайти на свои рабочие станции. В графической оснастке Windows — диспетчере сервера — стало невозможно подключиться к DNS (domain name system) с ошибкой «отказано в доступе». К DNS-серверу по командной строке тоже стало невозможно подключиться: `dnscmd/info` сообщал лишь «error access denied». В журнале стали появляться ошибки, что DNS пытается загрузить зону из контроллера домена, но операция завершалась неудачей. Диагностика контроллера домена выдала ошибку: `cdiag`. Системный администратор нашел ошибку: «не пройдена проверка `NCSecDesc`». Ошибка состояла в том, что контроллер домена предприятия (Active Directory) не работал. Это означало, что перенос данных со старого контроллера домена на новый, несмотря на сообщения операционной системы, что все прошло нормально, на самом деле не был завершён, а конфигурация контроллера домена повреждена. После всех попыток исправить конфигурацию в журнале стала появляться ошибка «8430 Внутренняя ошибка службы каталогов. Внутренний идентификатор: 3200db0». К проблеме с запуском DNS-сервера добавилась еще одна: клиентские персональные компьютеры не могли подключиться к контроллеру домена. Системному администратору удалось найти техническую статью на

сайте Microsoft о том, как можно исправить подключение клиентов. В ней было указано, что с помощью графической оснастки невозможно осуществить сброс пароля для учетной записи персонального компьютера. Вместо этого необходимо использовать командную строку и встроенную утилиту Netdom.exe. С помощью этой утилиты удалось провести сброс пароля для каждого пользователя, после чего они смогли работать. После перезагрузки сервера и повторных проверок ошибки пропали. Но на следующий день возникла другая проблема: пользователи не могли зайти на сервер терминалов. Появлялось сообщение о том, что домен не доступен. После проверки выяснилось, что отсутствовали папки с общим доступом, обеспечивающие подключение пользователей к серверу: netlogon, sysvol. Для исправления этой ошибки необходимо было выполнить ряд действий, описанных в технической статье Microsoft. После этого пользователи смогли работать с сервером нормально.

Пример 4. Рассмотрим пример порождения СД с помощью данных, используемых в ИТ. В работе [6] рассмотрены ИТ, представимые в виде DAG (Directed Acyclic Graph). Прописными латинскими буквами A, B, \dots будем обозначать *данные* (объекты), служащие входными или выходными данными преобразований информации в ИТ, на рисунке данные будем представлять окружностями. Преобразования будем называть *блоками*, обозначать строчными латинскими буквами и на рисунке обозначать прямоугольниками.



Векторные входные данные

Каждый блок соответствует преобразованию информации и решает одну или несколько задач, необходимых для реализации ИТ. Дуги DAG соответствуют передаче данных блокам от предыдущих блоков, т. е. дуга идет из выходных данных выполненного преобразования во входные данные для следующего преобразования.

Представляет интерес следующая схема (см. рисунок).

Пусть в ИТ, представленной в виде DAG, подаются входные данные X_1, \dots, X_k на преобразования g_1, \dots, g_k . Обозначим через $J(f)$ всевозможные выходные данные функции f . Тогда

$$J(f) = J^+(f) \cup J^-(f), \quad J^+(f) \cap J^-(f) = \phi,$$

где $J^+(f)$ — допустимые результаты f , $J^-(f)$ — недопустимые результаты, включая останов выполнения преобразования f .

Положим $D^+(f) = f^{-1}(J^+(f))$ — допустимая область определения f , $D^-(f) = f^{-1}(J^-(f))$ — область порождения недопустимых значений f . Предположим, что

$$B = f(A) = f_2(f_1(A)).$$

Тогда если функции f_1 и f_2 вычисляются правильно и $f(A) \in J^+(f)$, то $A \in D^+(f_2)$.

Рассмотрим схему на рисунке. Если результат выполнения функции f_1 принадлежит $J^+(f_1)$, а результат выполнения функции f_2 принадлежит $J^+(f_2)$, то отсюда не следует, что входные данные функции f принадлежат множеству $D^+(f)$. В самом деле, если $J^+(f_1) \times J^+(f_2) \not\subset D^+(f)$, то существуют $A \in J^+(f_1)$ и $B \in J^+(f_2)$ такие, что вектор $(A, B) \in D^-(f)$, т. е. $f(A, B)$ принимает недопустимые значения. Это значит, что если X_i принадлежит допустимой области определения функции, для которой она служит входной переменной, и существует хотя бы одна схема типа представленной на рисунке, такая что $J^+(f_1) \times J^+(f_2) \not\subset D^+(f)$, то существует набор значений X_1, \dots, X_k , при которых значение результата ИТ будет недопустимым.

Эту ситуацию можно рассматривать как СД, т. е. если противник вычислил условия, при которых все правильные исходные данные порождают ошибку, то он может извне нанести ущерб выполнению ИТ.

4 Теория распространения ошибок

Ущерб от сбоя в одном экземпляре ИТ не всегда велик, даже с учетом необходимости поиска причины сбоя. Поэтому эффективность использования противником СД для нанесения ущерба существенно зависит от влияния ошибок, порожденных с помощью СД. Очевидно, что влияние зависит от взаимодействия измененных значений параметра с другими параметрами системы.

Наибольшее значение в распространении ошибки имеет изменение в конфигурационных параметрах программного обеспечения. Во-первых, такие ошибки трудно находить, а во-вторых, конфигурационные ошибки повторяются при всех вычислениях, использующих этот параметр. Распространение ошибки рассмотрим на следующем примере.

Пример 5. Спустя 2 года после событий, описанных в примере 3, в организации вновь возникли проблемы с контроллером домена. Но в этот раз ошибка была сразу идентифицирована. С помощью утилиты `dcdiag` было получено сообщение «не пройдена проверка NetLogons». Восстановить необходимые сетевые ресурсы для подключения пользователей с помощью указанной ранее технической статьи Microsoft не удавалось — ошибка все равно присутствовала. Это означало, что причина ошибки на этот раз находилась в другом месте. С помощью утилиты `ntdsutil` удалось провести проверку базы данных контроллера домена. Было получено сообщение «операция прекращена с ошибкой — 1206 (JET_errDatabaseCorrupted, Non database file or corrupted db)». С помощью этой утилиты не удалось восстановить базу данных, поэтому необходимо было использовать другую утилиту: `esentutl`. Данная утилита в режиме «исправления» позволила устранить сбой в базе данных `esentutl/p C:\Windows\NTDS\ntds.dit`. Результатом выполнения стало сообщение «Integrity check successful. Operation completed successfully

in 2.308 seconds». Проведенное тестирование всей файловой системы, а также жесткого диска не выявило никаких проблем. Анализ показал, что данный сбой был вызван накопившимися ошибками двухлетней давности. Не исключено, что эта ошибка могла бы проявиться раньше или не проявиться вообще. Важно отметить, что резервный контроллер домена одновременно с основным перестал осуществлять подключение клиентов и сообщал об аналогичных ошибках. Это было связано с тем, что из-за поврежденной базы данных перестала работать репликация данных между серверами.

Проблеме распространения ошибок посвящено много исследований (см., например, [7–11]). В основе этих исследований лежит множество внутренних зависимостей различных сущностей компьютерных систем и сетей. Помимо очевидных, явных зависимостей существуют скрытые зависимости, выявление которых представляет значительные трудности. Все исследования в этой области ведутся методом моделирования.

В большинстве работ рассматриваются ошибки во взаимодействующих системах. Здесь получены интересные результаты [12], связанные со временем распространения сбоя, в реальной системе составившим 6–9 ч.

Как отмечается в работе [13], наиболее сложные задачи связаны с возможностью возникновения очередей, при этом очереди часто порождают рассинхронизацию технологических или информационных процессов.

5 Поиск BackDoors

Стандартные методы поиска вредоносного кода не работают при поиске СД. В самом деле, наиболее эффективным методом поиска уязвимостей, связанных с вредоносным кодом, считается интеллектуальный фаззинг. Этот метод позволяет искать сбои в преобразованиях путем слабой модификации данных, которые не дают сбоев.

Пример 4 СД с помощью подбора правильных данных, таких что сбой происходит на преобразованиях типа представленных на рисунке, показывает, что для создания такого СД необходимо исследовать области определения и значений преобразования с векторным входом во всей ИТ. Для проведения такого анализа необходимо эффективно вычислять функции принадлежности указанных множеств.

Одновременно можно показать, что указанный тип СД невозможен, если для всех схем типа представленных на рисунке (с поправкой на число входных переменных) выполняется условие:

$$J^+(f_1) \times J^+(f_2) \subseteq D^+(f).$$

В общем случае анализ угроз, порождаемых СД, надо начинать с исследований взаимозависимостей переменных и влияния таких параметров на работу

системы. Такой предварительный анализ значительно сократит число потенциальных угроз. Далее необходимо понять, какие значения выявленных параметров следует считать критическими для нанесения максимального ущерба.

Описанная подготовка к поиску СД является разумной, но не учитывает возможность существования и одновременной реализации нескольких СД (см. пример 4). Но даже в случае отдельного параметра требуется рассчитывать, насколько полно проведен поиск.

Как и в случае поиска уязвимостей в компьютерной системе, следует искать пути достижения выявленных критических параметров. В работе [12] утверждается, что решение этой задачи возможно только при точной имитирующей модели реальной системы.

Математическая модель распространения сбоя рассматривалась, например, в работе [14]. В этой работе построен точный алгоритм решения этой задачи, но он оказался NP-трудным. Построенный в этой же работе эвристический алгоритм работает быстрее, но примерно на 6% хуже, чем точный. Это значит, что появляется возможность пропуска распространения критической ошибки с помощью СД.

Для построения сложного СД можно использовать цепочки причинно-следственных связей, которые конструируются с помощью интеллектуального анализа данных.

6 Заключение

В работе отражен ряд проблем, связанных с угрозами распределенным ИС из-за наличия СД. Важно отметить, что большинство проблем, порожденных СД, приходится решать системным администраторам.

BackDoors могут порождаться недостаточностью (скрываемостью) информации о путях изменения конфигураций программного обеспечения, как это показано в примере 3. А в примере 1 системный администратор вообще не мог получить доступ к среде, в которой находится СД.

Наибольшую опасность представляют СД, порождаемые данными (см. пример 4). Это связано с тем, что построение таких СД и их анализ основаны на решении сложных математических задач, что не всегда доступно системному администратору.

Размер ущерба от существования СД связан с возможностью распространения ошибки при изменении значения параметра с помощью СД. Пример 5 показывает, что ошибка может не только распространяться в системе при воздействии на параметр через СД, но и накапливаться неопределенное время, что может породить критические проблемы.

В работе показано, что существующие технологии поиска уязвимостей в программно-аппаратной системе не решают проблемы поиска СД. Математические модели и методы решения задач поиска СД, кроме частных случаев, еще не

разработаны. Принципиально, что СД не связаны с внедрением в компьютерную систему вредоносного кода и с его использованием для нанесения ущерба. Поэтому, как правило, механизм нанесения ущерба через СД отличается от действия вредоносного кода.

Литература

1. *Тимонина Е. Е.* Анализ угроз скрытых каналов и методы построения гарантированно защищенных распределенных автоматизированных систем: Дис. . . . д-ра техн. наук. — М., 2004. 204 с.
2. *Грушо А. А., Грушо Н. А., Тимонина Е. Е.* Методы защиты информации от атак с помощью скрытых каналов и враждебных программно-аппаратных агентов в распределенных системах // Вестник РГГУ. Сер.: Документоведение и архивоведение. Информатика. Защита информации и информационная безопасность, 2009. № 10. С. 33–45.
3. *Грушо А. А., Забейайло М. И., Зацаринный А. А., Николаев А. В., Писковский В. О., Тимонина Е. Е.* Классификация ошибочных состояний в распределенных вычислительных системах и источники их возникновения // Системы и средства информатики, 2017. Т. 27. № 2. С. 29–40.
4. *Грушо А. А., Забейайло М. И., Зацаринный А. А., Николаев А. В., Писковский В. О., Сенчило В. В., Сударииков И. В., Тимонина Е. Е.* Об анализе ошибочных состояний в распределенных вычислительных системах // Системы и средства информатики, 2018. Т. 28. № 1. С. 99–109.
5. *Skorobogatov S., Woods C.* Breakthrough silicon scanning discovers backdoor in military chip // Cryptographic hardware and embedded systems / Eds. E. Prouff, P. Schaumont. — Lecture notes in computer science ser. — Heidelberg: Springer, 2012. Vol. 7428. P. 23–40. doi: 10.1007/978-3-642-33027-8_2.
6. *Грушо Н. А., Грушо А. А., Забейайло М. И., Тимонина Е. Е.* Методы нахождения причин сбоев в информационных технологиях с помощью метаданных // Информатика и её применения, 2020. Т. 14. Вып. 2. С. 33–39. doi: 10.14357/19922264200205.
7. *Rinaldi S. M., Peerenboom J. P., Kelly T.* Identifying, understanding, and analyzing critical infrastructure interdependencies // IEEE Contr. Syst. Mag., 2001. Vol. 21. Iss. 6. P. 11–25.
8. *Eusgeld I., Kröger W.* Towards a framework for vulnerability analysis of interconnected infrastructures // 9th Probabilistic Safety Assessment & Management Conference (International) Proceedings, 2008. P. 107–116.
9. *Griot C.* Modelling and simulation for critical infrastructure interdependency assessment: A meta-review for model characterisation // Int. J. Critical Infrastructures, 2010. Vol. 6. Iss. 4. P. 363–379.
10. *Canzani E., Kaufmann H., Lechner U.* An operator-driven approach for modeling interdependencies in critical infrastructures based on critical services and sectors // Critical information infrastructures security / Eds. G. Havarneanu, R. Setola, H. Naspopoulos, S. Wolthusen. — Lecture notes in computer science ser. — Springer, 2017. Vol. 10242. P. 308–320. doi: 10.1007/978-3-319-71368-7_27.

11. *Di Pietro A., Porta L. L., Lavalle L., Pollino M., Rosato V., Tofani A.* Simulation of cascading outages in (inter)-dependent services and estimate of their societal consequences // *Critical information infrastructures security* / Eds. G. Havarneanu, R. Setola, H. Nassopoulos, S. Wolthusen. — *Lecture notes in computer science ser.* — Springer, 2017. Vol. 10242. P. 340–345. doi: 10.1007/978-3-319-71368-7_30.
12. *Nan C., Eusgeld I., Kröger W.* Hidden vulnerabilities due to Interdependencies between two systems // *Critical information infrastructures security* / Eds. B. M. Hämmerli, N. Kalstad Svendsen, J. Lopez. — *Lecture notes in computer science ser.* — Springer, 2013. Vol. 7722. P. 252–263. doi: 10.1007/978-3-642-41485-5_22.
13. *Esposito Amideo A., Scaparra M. P.* A synthesis of optimization approaches for tackling critical information infrastructure survivability // *Critical information infrastructures security* / Eds. G. Havarneanu, R. Setola, H. Nassopoulos, S. Wolthusen. — *Lecture notes in computer science ser.* — Springer, 2017. Vol. 10242. P. 75–87. doi: 10.1007/978-3-319-71368-7_7.
14. *Banerjee J., Sen A., Zhou C.* On auxiliary entity allocation problem in multi-layered interdependent critical infrastructures // *Critical information infrastructures security* / Eds. G. Havarneanu, R. Setola, H. Nassopoulos, S. Wolthusen. — *Lecture notes in computer science ser.* — Springer, 2017. Vol. 10242. P. 25–37. doi: 10.1007/978-3-319-71368-7_3.

Поступила в редакцию 15.03.21

HIDDEN IMPACT WITHOUT MALICIOUS CODE

A. A. Grusho¹, N. A. Grusho¹, M. I. Zabezhailo², and E. E. Timonina¹

¹Institute of Informatics Problems, Federal Research Center “Computer Sciences and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

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

Abstract: The paper reflects a number of problems related to threats to distributed information systems due to the presence of BackDoor. The usage of BackDoor to attack distributed computer systems has great advantages over the traditional usage of malicious code. In fact, the usage of BackDoor is difficult to identify as a special attack, i. e., distinguish from error in computing processes. However, building attacks with BackDoor is much more difficult than building attacks with malicious code. The example of BackDoor is built using the initial data of information technology. This approach to the construction of BackDoor was not found in any of the descriptions of BackDoor. The paper shows that the existing technologies for finding vulnerabilities in the hardware and software of the system do not solve the problems of searching for BackDoor.

Mathematical models and methods for solving BackDoor search problems, except in particular cases, have not yet been developed. It is important that BackDoor is not associated with the introduction of malicious code into the computer system and with the usage of malicious code to cause damage. Therefore, as a rule, the mechanism of causing damage through BackDoor differs from the action of malicious code.

Keywords: information security; BackDoor; BackDoor through data

DOI: 10.14357/08696527210201

Acknowledgments

The paper was partially supported by the Russian Foundation for Basic Research (project 18-29-03081).

References

1. Timonina, E. E. 2004. Analiz ugroz skrytykh kanalov i metody postroeniya garantirovanno zashchishchennykh raspredelennykh avtomatizirovannykh sistem [The analysis of threats of covert channels and methods of creation of guaranteed protected distributed automated systems]. Moscow: Russian State University for the Humanities. D.Sc. Diss. 204 p.
2. Grusho, A., N. Grusho, and E. Timonina. 2009. Metody zashchity informatsii ot atak s pomoshch'yu skrytykh kanalov i vrazhdebnykh programmno-apparatnykh agentov v raspredelennykh sistemakh [Methods of information protection against covert channels attacks and malicious software/hardware agents in distributed systems]. *Vestnik RGGU. Ser. Dokumentovedenie i arkhivovedenie. Informatika. Zashchita informatsii i informatsionnaya bezopasnost'* [RGGU Bulletin. Document science and archive science. Informatics. Information security and information security ser.] 10:33–45.
3. Grusho, A. A., M. I. Zabezhalo, A. A. Zatsarinny, A. V. Nikolaev, V. O. Piskovski, V. V. Senchilo, and E. E. Timonina. 2017. Klassifikatsiya oshibochnykh sostoyaniy v raspredelennykh vychislitel'nykh sistemakh i istochniki ikh vozniknoveniya [Erroneous states classification in distributed computing systems and sources of their occurrence]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 27(2):29–40.
4. Grusho, A. A., M. I. Zabezhalo, A. A. Zatsarinny, A. V. Nikolaev, V. O. Piskovski, V. V. Senchilo, I. V. Sudarikov, and E. E. Timonina. 2018. Ob analize oshibochnykh sostoyaniy v raspredelennykh vychislitel'nykh sistemakh [About the analysis of erratic statuses in the distributed computing systems]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(1):99–109.
5. Skorobogatov, S., and C. Woods. 2012. Breakthrough silicon scanning discovers backdoor in military chip. *Cryptographic hardware and embedded systems*. Eds. E. Prouff and P. Schaumont. Lecture notes in computer science ser. Heidelberg: Springer. 7428:23–40. doi: 10.1007/978-3-642-33027-8_2.
6. Grusho, N. A., A. A. Grusho, M. I. Zabezhalo, and E. E. Timonina. 2020. Metody nakhozhdeniya prichin sboev v informatsionnykh tekhnologiyakh s pomoshch'yu metadannykh [Methods of finding the causes of information technology failures by

- means of meta data]. *Informatika i ee Primeneniya — Inform. Appl.* 14(2):33–39. doi: 10.14357/19922264200205.
7. Rinaldi, S. M., J. P. Peerenboom, and T. Kelly. 2001. Identifying, understanding, and analyzing critical infrastructure interdependencies. *IEEE Contr. Syst. Mag.* 21(6):11–25.
 8. Eusgeld, I., and W. Kröger. 2008. Towards a framework for vulnerability analysis of interconnected infrastructures. *9th Probabilistic Safety Assessment & Management Conference (International) Proceedings.* 107–116.
 9. Griot, C. 2010. Modelling and simulation for critical infrastructure interdependency assessment: A meta-review for model characterisation. *Int. J. Critical Infrastructures* 6(4):363–379.
 10. Canzani, E., H. Kaufmann, and U. Lechner. 2017. An operator-driven approach for modeling interdependencies in critical infrastructures based on critical services and sectors. *Critical information infrastructures security.* Eds. G. Havarneanu, R. Setola, H. Nassopoulos, and S. Wolthusen. Lecture notes in computer science ser. Springer. 10242:308–320. doi: 10.1007/978-3-319-71368-7_27.
 11. Di Pietro, A., L. L. Porta, L. Lavallo, M. Pollino, V. Rosato, and A. Tofani. 2017. Simulation of cascading outages in (inter)-dependent services and estimate of their societal consequences. *Critical information infrastructures security.* Eds. G. Havarneanu, R. Setola, H. Nassopoulos, and S. Wolthusen. Lecture notes in computer science ser. Springer. 10242:340–345. doi: 10.1007/978-3-319-71368-7_30.
 12. Nan, C., I. Eusgeld, and W. Kröger. 2013. Hidden vulnerabilities due to interdependencies between two systems. *Critical information infrastructures security.* Eds. B. M. Hämmerli, N. Kalstad Svendsen, and J. Lopez. Lecture notes in computer science ser. Springer. 7722:252–263. doi: 10.1007/978-3-642-41485-5_22.
 13. Esposito Amideo, A., and M. P. Scaparra. 2017. A synthesis of optimization approaches for tackling critical information infrastructure survivability. *Critical information infrastructures security.* Eds. G. Havarneanu, R. Setola, H. Nassopoulos, and S. Wolthusen. Lecture notes in computer science ser. Springer. 10242:75–87. doi: 10.1007/978-3-319-71368-7_7.
 14. Banerjee, J., A. Sen, and C. Zhou. 2017. On auxiliary entity allocation problem in multi-layered interdependent critical infrastructures. *Critical information infrastructures security.* Eds. G. Havarneanu, R. Setola, H. Nassopoulos, and S. Wolthusen. Lecture notes in computer science ser. Springer. 10242:25–37. doi: 10.1007/978-3-319-71368-7_3.

Received March 15, 2021

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

Grusho Nikolai A. (b. 1982) — Candidate of Science (PhD) in physics and mathematics, senior 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; info@itake.ru

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

АНАЛИТИЧЕСКАЯ МОДЕЛЬ ПРОТОКОЛА МНОЖЕСТВЕННОГО ДОСТУПА С ПРОСЛУШИВАНИЕМ КАНАЛА ДЛЯ ПРИЛОЖЕНИЙ ИНДУСТРИАЛЬНОГО ИНТЕРНЕТА ВЕЩЕЙ*

А. С. Царев¹, Э. М. Хайров², Ю. В. Гайдамака³, С. Я. Шоргин⁴

Аннотация: Индустриальный интернет вещей стал одной из ключевых концепций четвертой промышленной революции — нового подхода к построению промышленных процессов с использованием новейших технологий. Однако для планирования и выстраивания новых автоматизированных процессов необходимо понимать характеристики и ограничения, возникающие из-за технических особенностей системы. Для решения этой задачи в одном из сценариев использования промышленного интернета вещей предлагается математическая модель протокола многостанционного доступа с контролем несущей и предотвращением конфликтов. Такая модель позволяет оценить среднее время цикла передачи данных с датчика, вероятность успешной передачи данных, вероятность коллизии при передаче данных. Получены аналитические выражения для вычисления ключевых метрик производительности на уровне управления доступом к среде.

Ключевые слова: индустриальный интернет вещей; управление доступом к среде; прослушивание канала; множественный доступ; цепь Маркова; вероятность коллизии

DOI: 10.14357/08696527210202

1 Введение

Достижения технологического прогресса, такие как интернет вещей, расширяют возможности для улучшения производственных процессов. Индустриальный интернет вещей характеризуется миниатюризацией приемопередающих устройств, увеличенной плотностью устройств и более строгими ограничениями, предъявляемыми к задержкам [1]. Перечисленные требования могут быть

* Публикация выполнена при поддержке Программы стратегического академического лидерства РУДН и при финансовой поддержке РФФИ (проекты 19-07-00933 и 20-07-01064).

¹ Российский университет дружбы народов, atsarev@sci.pfu.edu.ru

² Российский университет дружбы народов, emil.khayrov@gmail.com

³ Российский университет дружбы народов; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, gaydamaka-yuv@rudn.ru

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

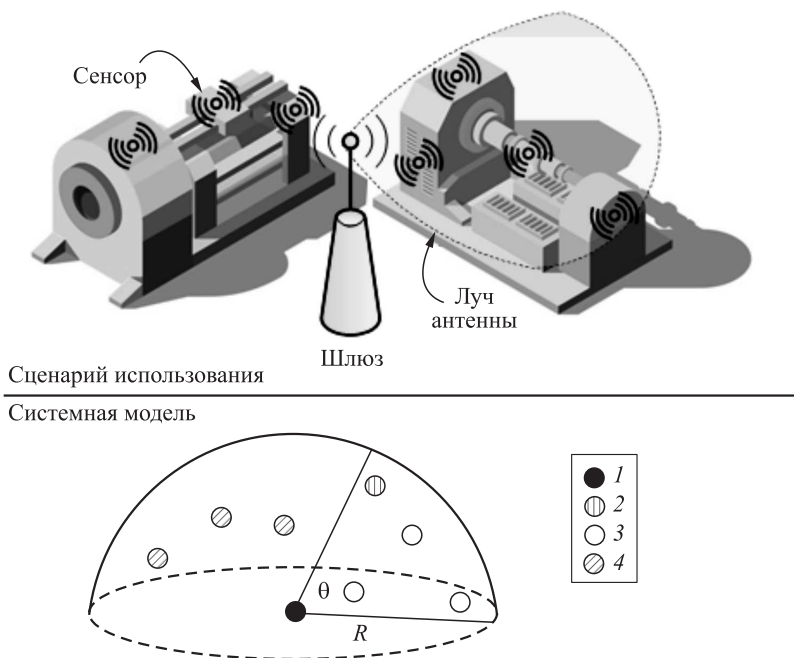


Рис. 1 Системная модель приложения промышленного интернета вещей: 1 — шлюз; 2 — передающий узел; 3 — узел в секторе покрытия антенны; 4 — узел вне сектора покрытия антенны

удовлетворены в существующих беспроводных сетях при условии передачи данных в миллиметровом / терагерцевом диапазоне волн. Высокий уровень потерь при распространении волн может быть компенсирован использованием направленных антенн [2, 3].

Рассмотрим сценарий использования промышленного интернета вещей, представленный на рис. 1.

Некоторое предприятие оборудовано промышленными станками, каждый из которых имеет некоторое число сенсоров, оснащенных передатчиками. Сенсоры считывают показания и передают их шлюзу по беспроводному каналу связи согласно протоколу управления доступом к среде [3]. Шлюз оснащен набором направленных антенн, меняющих направление сектора покрытия θ согласно расписанию. Каждый передатчик оснащен одной всенаправленной антенной, используемой для передачи данных через шлюз. Передача накопленных сенсорами показаний начинается в случайный момент времени в течение периода покрытия антенной шлюза соответствующего сектора пространства, что может привести к потере информации на шлюзе вследствие коллизий, возникающих при одновременных попытках передачи показаний от нескольких сенсоров. Опи-

санная ситуация деградации пропускной способности системы актуальна и для миллиметрового диапазона волн [4].

2 Процедура множественного доступа с прослушиванием канала и предотвращением коллизий

Для предотвращения деградации производительности системы предлагается использовать модификацию протокола многостанционного доступа с контролем несущей и предотвращением конфликтов (Carrier Sense Multiple Access With Collision Avoidance, CSMA/CA) [5] при помощи инициализации сессии передачи данных через RTS/CTS-процедуру (RTS/CTS — Request To Send / Clear To Send, запрос на отправку/разрешение отправки). В таком случае подразумевается, что возможны коллизии между сигнальными сообщениями, а коллизии между данными будут сведены к нулю. Учитывая, что в реальных приложениях размер пакета данных на порядок больше размера сигнального сообщения, такое улучшение позволит более эффективно использовать имеющийся канал.

На рис. 2 представлена временная диаграмма предлагаемого протокола для шлюза и двух узлов (передатчиков). При наличии данных перед их передачей разыгрывается равномерно распределенная отсрочка передачи. После этого узел отправляет RTS-сообщение. В случае его приема без коллизий шлюз проводит широковещательную рассылку CTS-сообщения, содержащего информацию об узле, который займет канал своими данными в следующий временной интервал. Узел, не получивший от шлюза разрешения на передачу данных до оконча-

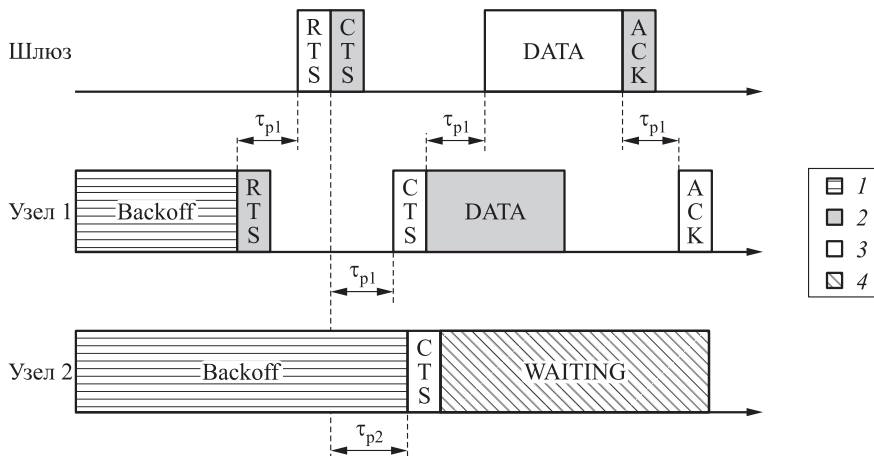


Рис. 2 Временная диаграмма, описывающая успешную передачу данных с помощью RTS/CTS-процедуры: 1 — отсрочка передачи; 2 — передача сообщения; 3 — прием сообщения; 4 — период «канал занят»; τ_p — задержка распространения

ния заданного тайм-аута, что возможно вследствие коллизии RTS-сообщений на шлюзе, начинает новую попытку передачи с увеличенной отсрочкой перед отправкой нового RTS-сообщения.

3 Математическая модель процедуры установления соединения передатчиком

Функционирование передатчика согласно процедуре множественного доступа с прослушиванием канала и предотвращением коллизий моделируется с помощью цепи Маркова в дискретном времени.

Пусть множество X описывает конечное пространство состояний передатчика:

$$X = \{ \text{idle}, \text{backoff}_1^1, \dots, \text{backoff}_{1 T_{\max}}^1, \text{wait}^1, \text{rts}^1, \text{out}^1, \dots, \text{cts}, \text{data}, \text{ack} \}.$$

Согласно [6], в каждый момент времени передатчик может быть в одном из перечисленных состояний. Состояние *idle* отвечает за слот времени, когда у сенсора нет данных для передачи, где вероятность передачи данных равна p_a . Отсрочка *backoff* перед отправкой поделена на соответствующее число состояний backoff_j^i , $1 \leq i \leq N + 1$, $1 \leq j \leq iT_{\max}$, где N — максимальное число допустимых попыток передачи RTS-сообщения; T_{\max} — максимальная длительность отсрочки для первой попытки. Из backoff_j^i передатчик может перейти в состояние backoff_{j-1}^i с вероятностью p_f или в wait^i , $1 \leq i \leq N + 1$, с вероятностью $1 - p_f$, где p_f — вероятность того, что в канале отсутствует CTS-сообщение. При передаче RTS-сообщения с вероятностью p может возникнуть коллизия между RTS-сообщениями. Также p учитывает случай, когда поступившее RTS-сообщение игнорируется шлюзом в связи с исходящей передачей CTS-сообщения.

Решив систему уравнений, соответствующую цепи Маркова на рис. 3, получим следующее распределение вероятностей:

$$\begin{aligned}
 p_{\text{idle}} + \sum_{i=1}^{N+1} (p_{\text{rts}}^i + p_{\text{out}}^i + p_{\text{backoff}}^i + p_{\text{wait}}^i) + p_{\text{cts}} + p_{\text{data}} + p_{\text{ack}} &= 1; \\
 p_{\text{rts}}^i &= p_a p^{i-1} p_{\text{idle}}; \\
 p_{\text{out}}^i &= p_a p^i p_{\text{idle}}; \\
 p_{\text{backoff}}^i &= \sum_{j=1}^{iT_{\max}} p_{\text{backoff}(j)}^i = \\
 &= p_a p^{i-1} p_{\text{idle}} \frac{iT_{\max} + 1}{iT_{\max}} \frac{p_f (p_f^{iT_{\max}} - iT_{\max} - 1) + iT_{\max}}{(1 - p_f^{iT_{\max}+1})(1 - p_f)};
 \end{aligned}$$

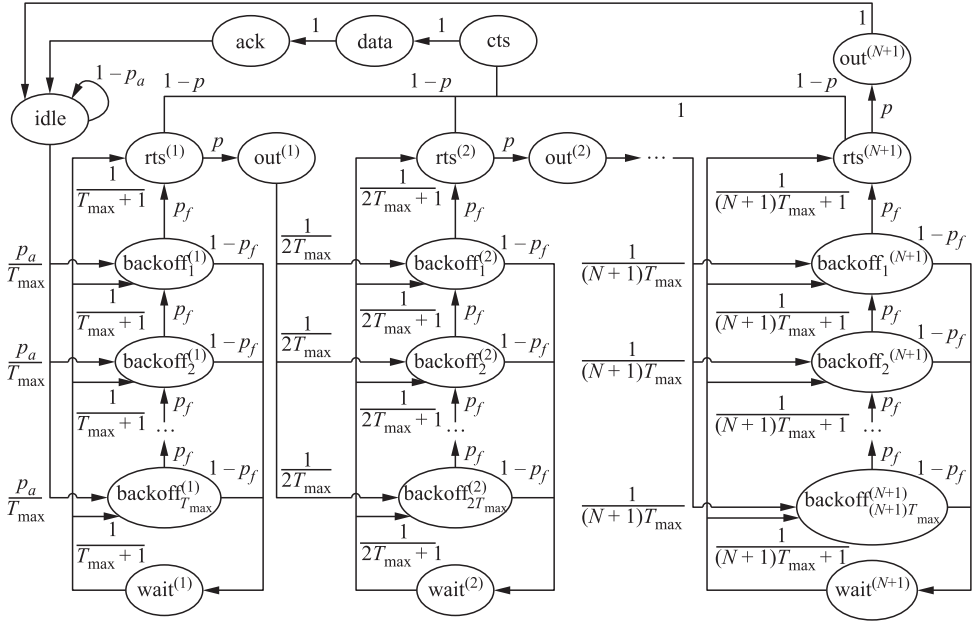


Рис. 3 Цепь Маркова с возможными состояниями передатчика

$$p_{\text{wait}}^i = p_a p^{i-1} p_{\text{idle}} \frac{i T_{\text{max}} + 1}{i T_{\text{max}}} \frac{p_f (p_f^{i T_{\text{max}}} - i T_{\text{max}} - 1) + i T_{\text{max}}}{1 - p_f^{i T_{\text{max}} + 1}};$$

$$p_{\text{cts}} = p_{\text{data}} = p_{\text{ack}} = \sum_{i=1}^{N+1} p_a p^{i-1} (1-p) p_{\text{idle}}.$$

Пусть $\chi(t)$ — длительность цикла передачи данных (независимо от того, закончился ли он успешной передачей данных), притом T_s , $s \in \{\text{idle}, \text{rts}, \text{out}, \text{backoff}, \text{wait}, \text{cts}, \text{data}, \text{ack}\}$, — длительность нахождения в соответствующем состоянии. Тогда среднюю длительность цикла передачи данных можно выразить следующим образом:

$$M[\chi] p_{\text{idle}} = \sum_{i=1}^{N+1} T_s p_s^i =$$

$$= p_{\text{idle}} T_{\text{idle}} + \sum_{i=1}^{N+1} p_a p^{i-1} p_{\text{idle}} \left(T_{\text{rts}} + p T_{\text{out}} + (1-p)(T_{\text{cts}} + T_{\text{data}} + T_{\text{ack}}) + \right.$$

$$\left. + \left(T_{\text{wait}} + \frac{T_{\text{backoff}}}{1-p_f} \right) \left(\frac{i T_{\text{max}} + 1}{i T_{\text{max}}} \right) \frac{p_f (p_f^{i T_{\text{max}}} - i T_{\text{max}} - 1) + i T_{\text{max}}}{1 - p_f^{i T_{\text{max}} + 1}} \right).$$

В таком случае стационарное распределение вероятностей с учетом средней продолжительности передачи данных будет иметь вид:

$$p_{\text{RTS}} = \frac{\sum_{i=1}^{N+1} p_a p^{i-1} T_{\text{rts}}}{M[\chi]} ;$$

$$p_{\text{OUT}} = \frac{\sum_{i=1}^{N+1} p_a p^i T_{\text{out}}}{M[\chi]} ;$$

$$P_{\text{BACKOFF}} =$$

$$= \frac{1}{M[\chi]} \sum_{i=1}^{N+1} p_a p^{i-1} \frac{i T_{\text{max}} + 1}{i T_{\text{max}}} \frac{p_f (p_f^{i T_{\text{max}}} - i T_{\text{max}} - 1) + i T_{\text{max}}}{(1 - p_f^{i T_{\text{max}+1}}) (1 - p_f)} T_{\text{backoff}} ;$$

$$p_{\text{WAIT}} = \frac{1}{M[\chi]} \sum_{i=1}^{N+1} p_a p^{i-1} \frac{i T_{\text{max}} + 1}{i T_{\text{max}}} \frac{p_f (p_f^{i T_{\text{max}}} - i T_{\text{max}} - 1) + i T_{\text{max}}}{1 - p_f^{i T_{\text{max}+1}}} T_{\text{wait}} ;$$

$$p_{\text{CTS}} = \frac{\sum_{i=1}^{N+1} p_a p^{i-1} (1 - p) T_{\text{cts}}}{M[\chi]} ;$$

$$p_{\text{DATA}} = \frac{\sum_{i=1}^{N+1} p_a p^{i-1} (1 - p) T_{\text{data}}}{M[\chi]} ;$$

$$p_{\text{ACK}} = \frac{\sum_{i=1}^{N+1} p_1 p^{i-1} (1 - p) T_{\text{ack}}}{M[\chi]} .$$

Сформулируем вероятностно-временные характеристики, позволяющие оценить производительность системы, функционирующей в режиме множественного доступа. Основную роль будут играть вероятность коллизии p , вероятность успешной передачи данных p_{success} и вероятность неуспешной передачи данных после p_{failure} . Кроме того, интерес представляет изучение средней длительности цикла передачи данных $M[\chi]$ [7].

Определим p_{success} в следующем виде:

$$p_{\text{success}} = \sum_{i=1}^{N+1} p_a p^{i-1} (1 - p) .$$

Тогда p_{failure} примет вид:

$$p_{\text{failure}} = p_a p^{N+1} = 1 - p_{\text{success}} .$$

Вероятность p вычисляется как решение трансцендентного уравнения [8]

$$p = 1 - (1 - p_{\text{RTS}})^{n_{\theta}-1} + 1 - (1 - p_{\text{CTS}})^{n_{\theta}-1} ;$$
$$p_f = (1 - p_{\text{CTS}})^{n_{\theta}-1} .$$

4 Заключение

В статье представлена математическая модель протокола множественного доступа с прослушиванием канала, учитывающая особенность приложений промышленного интернета вещей. Предложены метрики для оценки производительности системы на уровне доступа к среде, а также сформулированы аналитические выражения для нахождения вероятностно-временных характеристик, влияющих на пропускную способность.

В дальнейшем планируется провести комплексный анализ производительности системы, учитывая потери как на уровне доступа к среде (MAC layer), так и на физическом уровне (PHY layer) модели взаимодействия открытых систем.

Литература

1. *Aceto G., Persico V., Pescapé A.* A survey on information and communication technologies for industry 4.0: State-of-the-art, taxonomies, perspectives, and challenges // *IEEE Commun. Surv. Tut.*, 2019. Vol. 21. Iss. 4. P. 3467–3501. doi: 10.1109/COMST.2019.2938259.
2. *Elayan H., Amin O., Shubair R., Alouini M.* Terahertz communication: The opportunities of wireless technology beyond 5G // *Conference (International) on Advanced Communication Technologies and Networking Proceedings*. — Piscataway, NJ, USA: IEEE, 2018. Art ID: 8360286. 5 p. doi: 10.1109/COMMNET.2018.8360286.
3. *Buratti C., Mesini L., Verdone R.* Comparing MAC protocols for industrial IoT using Terahertz communications // *31st Annual Symposium (International) on Personal, Indoor and Mobile Radio Communications Proceedings*. — Piscataway, NJ, USA: IEEE, 2020. Art ID: 9217308. 7 p. doi: 10.1109/pimrc48278.2020.9217308.
4. *Чухно Н. В., Чухно О. В., Гайдамака Ю. В., Самуйлов К. Е.* Об условиях взаимного обнаружения устройств при передаче в миллиметровом диапазоне длин волн // *Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем: Сб. мат-лов VIII Всеросс. конф. (с международным участием) в рамках 54-й Всеросс. конф. по проблемам математики, информатики, физики и химии*. — М.: РУДН, 2018. С. 187–190.
5. Characteristics of broadband radio local area networks: Recommendation ITU-R M.1450-5, 2014.
6. *Samouylov K. E., Gaidamaka Yu. V., Gudkova I. A., Zaripova E. R., Shorgin S. Ya.* Baseline analytical model for machine-type communications over 3GPP RACH in LTE-advanced networks // *Computer and information sciences* / Eds. T. Czachórski, E. Gelembé, K. Grochla, R. Lent. — Communications in computer and information science book ser. — Springer, 2016. Vol. 659. P. 203–213. doi: 10.1007/978-3-319-47217-1_22.

7. *Tsarev A., Khayrov E., Medvedeva E., Gaidamaka Y., Buratti C.* Analytical model for CSMA-based MAC protocol for industrial IoT applications // Internet of Things, smart spaces, and next generation networks and systems / Eds. O. Galinina, S. Andreev, S. Balandin, Y. Koucheryavy. — Lecture notes in computer science. — Springer, 2020. Vol. 12526. P. 240–258. doi: 10.1007/978-3-030-65729-1_21.
8. *Bianchi G.* IEEE 802.11-saturation throughput analysis // IEEE Commun. Lett., 1998. Vol. 2. Iss. 12. P. 318–320. doi: 10.1109/4234.736171.

Поступила в редакцию 15.03.21

ANALYTICAL MODEL OF CARRIER SENSE MULTIPLE ACCESS PROTOCOL FOR INDUSTRIAL INTERNET OF THINGS APPLICATIONS

A. S. Tsarev¹, E. M. Khayrov¹, Yu. V. Gaidamaka^{1,2}, and S. Ya. Shorgin²

¹Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, 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: Industrial Internet of Things (IIoT) is one of the key features of Industry 4.0 that stands for a new approach to design industrial processes using the latest technological advances. However, before implementing new processes, it is necessary to understand the characteristics and limitations that arise due to technical complexity of the system. To solve this problem for some generalized IIoT use case, a mathematical model of the multiple access protocol with channel sensing is proposed. This model allows one to estimate the average sensor's data transmission cycle time, successful data transmission probability, and probability of a collision during data transmission. Analytical expressions for calculating key performance metrics at the access layer are proposed.

Keywords: Industrial Internet of Things (IIoT); MAC layer; channel sensing; multiple access; Markov chain; collision probability

DOI: 10.14357/08696527210202

Acknowledgments

The paper has been supported by the RUDN University Strategic Academic Leadership Program and funded by the Russian Foundation for Basic Research according to the research projects No. 19-07-00933 and No. 20-07-01064.

References

1. Aceto, G., V. Persico, and A. Pescapé. 2019. A survey on information and communication technologies for industry 4.0: State-of-the-art, taxonomies, perspectives, and challenges. *IEEE Commun. Surv. Tut.* 21(4):3467–3501. doi: 10.1109/COMST.2019.2938259.
2. Elayan, H., O. Amin, R. Shubair, and M. Alouini. 2018. Terahertz communication: The opportunities of wireless technology beyond 5G. *Conference (International) on Advanced Communication Technologies and Networking Proceedings*. Piscataway, NJ: IEEE. Art ID: 8360286. 5 p. doi: 10.1109/COMMNET.2018.8360286.
3. Buratti, C., L. Mesini, and R. Verdone. 2020. Comparing MAC protocols for industrial IoT using terahertz communications. *31st Annual Symposium (International) on Personal, Indoor and Mobile Radio Communications Proceedings*. Piscataway, NJ: IEEE. Art ID: 9217308. 7 p. doi: 10.1109/pimrc48278.2020.9217308.
4. Chukhno, N. V., O. V. Chukhno, Yu. V. Gaydamaka, and K. E. Samuylov. 2018. Ob usloviyakh vzaimnogo obnaruzheniya ustroystv pri peredache v millimetrovom diapazone dlin voln [MMWAVE channel access and devices' detection during data transmission in unlicensed bands]. *Informatsionno-telekommunikatsionnye tekhnologii i matematicheskoe modelirovanie vysokotekhnologichnykh sistem: Sb. mat-ov VIII Vseross. konf. (s mezhdunarodnym uchastiyem) v ramkakh 54-y Vseross. konf. po problemam matematiki, informatiki, fiziki i khimii* [8th All-Russian Conference (International) in the framework of the 54th All-Russian Conference on Mathematics, Informatics, Physics and Chemistry "Information and telecommunication technologies and mathematical modeling of high-tech systems" Proceedings]. Moscow: RUDN. 187–190.
5. Recommendation ITU-R M.1450-5. 2014. Characteristics of broadband radio local area networks. Available at: <https://www.itu.int/rec/R-REC-M.1450-5-201404-I> (accessed April 7, 2021).
6. Samouylov, K. E., Y. V. Gaidamaka, I. A. Gudkova, E. R. Zaripova, and S. Y. Shorgin. 2016. Baseline analytical model for machine-type communications over 3GPP RACH in LTE-advanced networks. *Computer and information sciences*. Eds. T. Czachórski, E. Gelenbe, K. Grochla, and R. Lent. Communications in computer and information science book ser. Springer. 659:203–213. 10.1007/978-3-319-47217-1_22.
7. Tsarev, A., E. Khayrov, E. Medvedeva, Y. Gaidamaka, and C. Buratti. 2020. Analytical model for CSMA-based MAC protocol for industrial IoT applications. *Internet of Things, smart spaces, and next generation networks and systems*. Eds. O. Galinina, S. Andreev, S. Balandin, and Y. Koucheryavy. Lecture notes in computer science ser. Springer. 12526:240–258. doi: 10.1007/978-3-030-65729-1_21.
8. Bianchi, G. 1998. IEEE 802.11-saturation throughput analysis. *IEEE Commun. Lett.* 2(12):318–320. doi: 10.1109/4234.736171.

Received March 15, 2021

Contributors

Tsarev Alexey S. (b. 1994) — PhD student, Department of Applied Probability and Informatics, Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; atsarev@sci.pfu.edu.ru

Khayrov Emil M. (b. 1997) — master student, Department of Applied Probability and Informatics, Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; emil.khayrov@gmail.com

Gaidamaka Yuliya V. (b. 1971) — Doctor of Science in physics and mathematics, professor, Department of Applied Probability and Informatics, Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; 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; gaydamaka-yuv@rudn.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

ОБЕСПЕЧЕНИЕ НАДЕЖНОСТИ ЦЕНТРА КОЛЛЕКТИВНОГО ПОЛЬЗОВАНИЯ ФИЦ ИУ РАН*

А. А. Зацаринный¹, А. И. Гаранин², С. А. Денисов³

Аннотация: Проведена оценка надежности центра коллективного пользования (ЦКП) ФИЦ ИУ РАН с учетом показателей надежности составляющих его систем: гибридного высокопроизводительного вычислительного комплекса (ГВВК), систем электропитания и кондиционирования. Проведен анализ путей повышения надежности ЦКП. Сформулированы основные условия обеспечения надежности функционирования и поддержания работоспособности аппаратных и аппаратно-программных средств (АПС) ЦКП в период их эксплуатации. Даны рекомендации по обеспечению качественного проведения работ по техническому обслуживанию и ремонту оборудования.

Ключевые слова: центр коллективного пользования; система электропитания; система кондиционирования; надежность; отказ; эквивалентная схема для расчета надежности

DOI: 10.14357/08696527210203

1 Введение

В [1] проведена оценка надежности ГВВК ФИЦ ИУ РАН, который создан на базе центра обработки данных (ЦОД) с использованием серверов архитектуры POWER различных модификаций (серверы Power 9 укомплектованы четырьмя графическими ускорителями Tesla V100, а также имеются серверы, оснащенные только классическими процессорами CPU).

Технические средства ЦКП расположены в двух зданиях. Средствами ЦОД организовано сетевое взаимодействие компонентов ГВВК между собой, с пользователями системы и вспомогательными сервисами, предоставляемыми облачной инфраструктурой ЦОД ФИЦ ИУ РАН.

По результатам расчетов надежность ГВВК с учетом показателей надежности автоматизированных рабочих мест пользователей, средств защиты информации и структурированной кабельной системы составила $K_T = 0,999922 \dots 0,999928$.

* Работа выполнена при частичной поддержке РФФИ (проекты 18-29-03091, 18-29-03100).

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

² Федеральний исследовательский центр «Информатика и управление» Российской академии наук, algaranin@mail.ru

³ Федеральний исследовательский центр «Информатика и управление» Российской академии наук, SDenisov@frccsc.ru

Помимо ГВВК, рассмотренного в [1], в состав ЦКП ФИЦ ИУ РАН входят и обеспечивающие элементы: система электропитания и система кондиционирования, надежность которых при указанных выше расчетах не учитывалась.

Цель настоящей статьи — оценка надежности ЦКП с учетом надежности ГВВК (результаты получены в [1]) и обеспечивающих его элементов, а также разработка направлений обеспечения надежности ЦКП в условиях постоянной эксплуатации.

2 Оценка надежности центра коллективного пользования

2.1 Система электропитания

Электропитание ЦОД обеспечивается от шести трансформаторов (рис. 1). Трансформаторы 1 и 2 в здании 1 взаимно резервируют друг друга в режиме «горячего резерва». При пропадании электропитания на вводе основного трансформатора или его отказе переключение на резервный трансформатор осуществляется устройством автоматического ввода резерва (АВР).

Электропитание серверного оборудования ГВВК в обоих зданиях обеспечивается через источники бесперебойного питания (ИБП). Другое оборудование (кондиционеры, автоматизированные рабочие места и др.) через сеть электропитания подключается непосредственно к распределительным шкафам (РЩ).

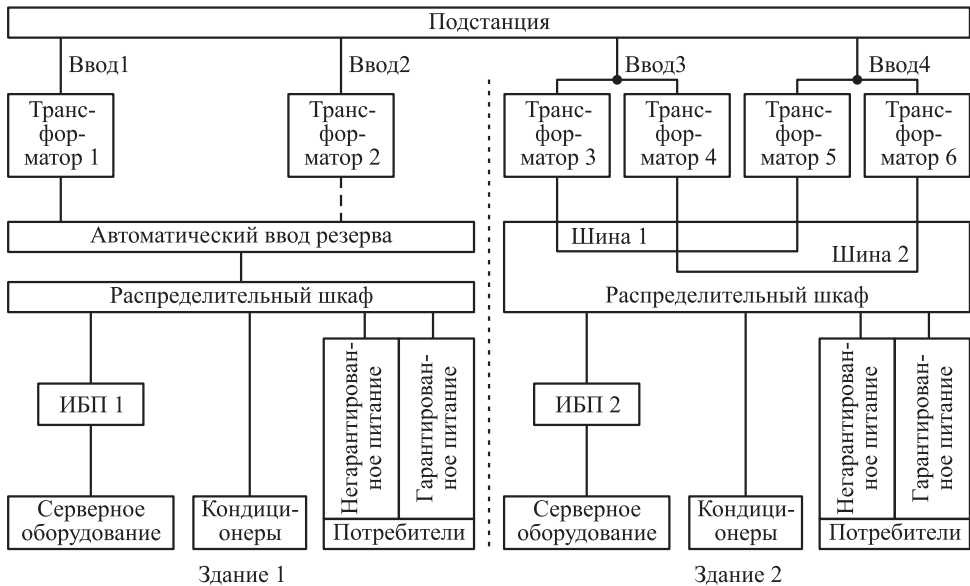


Рис. 1 Структурная схема системы электропитания ЦКП

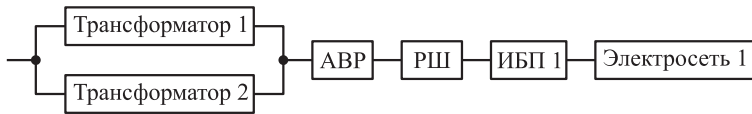


Рис. 2 Эквивалентная схема для расчета надежности системы электропитания в здании 1

В здании 2 электропитание обеспечивается от четырех трансформаторов. Трансформаторы 3 и 4 подключены к вводу 3, трансформаторы 5 и 6 — к вводу 4. В целях резервирования электропитания в случае выхода из строя ввода 3 или 4 трансформаторы 3 и 5 подключены к общей шине, трансформаторы 4 и 6 — к другой общей шине. К каждой из шин подключены свои группы пользователей.

Каждый из двух трансформаторов, подключенных к шине, способен в одиночку обеспечить нагрузку от подключенного к шине оборудования.

Оценим надежность представленной системы электропитания. Для этого сформулируем понятие ее «отказа».

Под **отказом** системы электропитания будем понимать событие, заключающееся в прекращении подачи электроэнергии на оборудование ГВВК и требующее для возобновления подачи электроэнергии проведения ремонтных работ с привлечением обслуживающего персонала.

Оценку надежности системы электропитания проведем с использованием эквивалентной схемы для расчета надежности [2]. Методику расчета рассмотрим на примере здания 1.

Эквивалентная схема для расчета надежности системы электропитания в здании 1 представлена на рис. 2.

Эта эквивалентная схема представляет собой последовательно-параллельное соединение элементов системы электропитания.

При *последовательном* соединении n элементов системы в случае экспоненциального распределения для средних времен безотказной работы системы (T_{oc}) и k -го элемента (T_{ok}) получаем [3, 4]:

$$T_{oc} = \left(\sum_{k=1}^n \frac{1}{T_{ok}} \right)^{-1}. \quad (1)$$

При *параллельном* соединении n элементов системы для случая, когда все элементы одинаковые ($T_{ik} = T_o$) и отказы элементов независимы между собой, а $T_o \gg T_b$, где T_b — время восстановления системы, для оценки среднего времени безотказной работы системы (T_{op}) можно воспользоваться выражением [3]:

$$T_{op} = T_o \sum_{k=1}^n \frac{1}{k} = T_o \left(1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} \right). \quad (2)$$

Таблица 1 Исходные данные по надежности элементов ЦКП

Устройство	T_o , ч	K_r
Трансформатор 1 и 2	260 000	0,999992
Трансформатор 3–6	400 000	0,999995
Устройство АВР	500 000	0,999996
ИБП 1	1 000 000	0,999998
ИБП 2	250 000	0,999992
Распределительный щит	500 000	0,999996
Сеть электропитания 1 и 2	400 000	0,999995
Ввод 1–4	300 000	0,999993
Кондиционер типа 1 (№№ 1–6)	22 500	0,999998
Кондиционер типа 2 (№№ 7–10)	40 000	0,999988

Коэффициент готовности оценим с использованием выражения [3]:

$$K_r = \frac{T_o}{T_o + T_b}. \quad (3)$$

Используя выражения (1)–(3) и данные из табл. 1, при $T_b = 2$ ч получим следующие количественные значения надежности подсистемы электропитания:

- в здании 1: $T_{oc1} = 99\,363$ ч; $K_{r1} = 0,999998$;
- в здании 2: $T_{oc2} = 84\,459$ ч; $K_{r2} = 0,999979$.

2.2 Система кондиционирования

В состав системы кондиционирования ЦОД (рис. 3) входят 10 кондиционеров. Функционирование серверного оборудования в здании 1 обеспечивают $\nu_1 = 6$ кондиционеров. Нормальную температуру в помещении обеспечивают $n_1 = 4$ кондиционера. Два кондиционера используются в качестве резервных и подключаются в случае повышения температуры в охлаждаемом помещении выше допустимого значения либо в случае выхода из строя или проведения профилактики на кондиционерах, используемых в качестве основных.

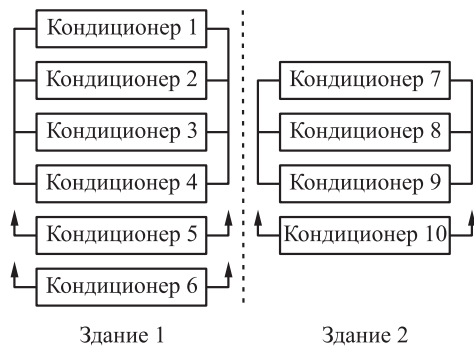


Рис. 3 Структурная схема системы кондиционирования ГВВК

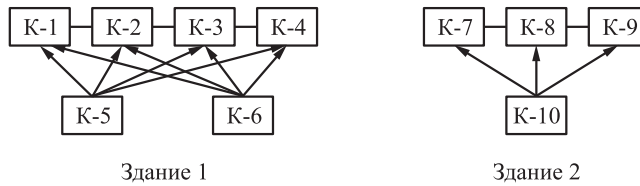


Рис. 4 Расчетно-логическая схема для расчета надежности системы кондиционирования

В здании 2 в состав системы входят $\nu_2 = 4$ кондиционера, три из которых используются в качестве основных ($n_2 = 3$) и один резервный.

Резервные кондиционеры подключаются автоматически при повышении температуры в охлаждаемом помещении выше установленного значения. Все кондиционеры взаимозаменяемы, поэтому в качестве резервных могут использоваться любые из установленных в зданиях кондиционеров.

Под **отказом** подсистемы кондиционирования будем понимать событие, заключающееся в снижении числа работоспособных кондиционеров в помещении ниже минимально необходимого и требующее для восстановления отказавших кондиционеров проведения ремонтных работ с привлечением обслуживающего персонала.

Для расчета надежности технической системы (ТС) при использовании структурного резервирования обычно составляется *расчетно-логическая схема* резервированной системы (рис. 4).

В теории надежности схемы, представленные на рис. 4, относятся к системам с ненагруженным резервом при резервировании с дробной кратностью, в которых отказ одного из n основных однотипных элементов приводит к включению на его место одного из $(\nu - n)$ резервных. При этом полагают, что элементы, находящиеся в резерве, отказывать не могут до их включения на место отказавшего основного элемента.

Исходя из этого условия и учитывая, что в процессе нормального функционирования ТС в работе находятся постоянно n элементов, средняя наработка до отказа (T_{oTC}) системы в целом за время t при экспоненциальном распределении может определяться из следующего выражения [3]:

$$T_{oTC} = T_o \frac{\nu - n + 1}{n}, \quad (4)$$

где T_o — средняя наработка на отказ основного и резервного элемента.

Коэффициент готовности системы оценим при $T_B = 0,5$ ч.

Используя расчетные выражения (3) и (4), оценим количественные значения надежности подсистемы кондиционирования с использованием исходных данных, представленных в табл. 1. Получим:

- в здании 1: $K_{rTC1} = 0,999997$; $T_{oTC1} = 16\,875$ ч;
- в здании 2: $K_{rTC2} = 0,999981$; $T_{oTC2} = 26\,667$ ч.

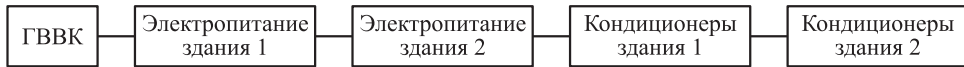


Рис. 5 Эквивалентная схема для расчета надежности ЦКП

Таблица 2 Результаты оценки надежности ЦКП и его элементов

Оцениваемая система	Здание	T_o , ч	K_T
ГВВК*		6 944	0,999922–0,999928
Вычислительная инфраструктура ГВВК*		131 578	0,9999962–0,9999977
Система электропитания	1	99 363	0,99998
	2	84 459	0,999976
Система кондиционирования	1	16 875	0,99997
	2	26 667	0,999981
Центр коллективного пользования		3 808	0,99987

* Оценка надежности проведена в [1].

Оценим показатели надежности *ЦКП в целом*.

Под **отказом** ЦКП будем понимать событие, заключающееся в прекращении ГВВК решения задач по причине неисправности аппаратно-программных средств ЦКП и требующее для восстановления процесса функционирования проведения ремонтных работ с привлечением обслуживающего персонала.

С учетом принятого понятия «отказа» составим эквивалентную схему для расчета надежности центра коллективного пользования (рис. 5).

Используя выражения (2) и (3), оценим надежность представленной схемы. Получим: $T_{oЦКП} = 3808$ ч; $K_{TЦКП} = 0,99987$.

Результаты проведенных выше расчетов отражены в табл. 2.

Анализ результатов, представленных в табл. 2, позволяет сделать вывод, что показатели надежности ГВВК (K_T) и систем, обеспечивающих его функционирование (электропитания и кондиционирования), составляют величины одного порядка, что отвечает требованиям по обеспечению равнопрочности элементов ЦКП.

3 Обеспечение надежности центра коллективного пользования в условиях постоянной эксплуатации

Обобщение и анализ опыта эксплуатации систем, аналогичных ЦКП, показывает, что основными условиями обеспечения надежности функционирования и поддержания работоспособности аппаратных и аппаратно-программных средств в составе современных сложных ТС, созданных на основе наукоемких технологий, в период их эксплуатации являются:

1. Качественная организация эксплуатации АПС, в ходе которой должны быть организованы и проводиться следующие мероприятия:
 - выполнение установленных в технических описаниях и инструкциях по эксплуатации условий и правил применения технических и программных средств;
 - формирование и поддержание рационального комплекта запасных частей, инструментов и принадлежностей, оптимизация его состава по номенклатуре и количеству;
 - использование измерительной и проверочной аппаратуры и программных средств диагностики технических и программных средств;
 - поддержание требуемой квалификации пользователей, администраторов и обслуживающего персонала ЦКП за счет обучения по специальности, самостоятельного изучения технической и эксплуатационной документации на объекте.
2. Своевременное и качественное проведение технического обслуживания, включая диагностику технических средств в целях раннего выявления причин отказов и принятия мер по их устранению.
3. Обеспечение внешних условий функционирования ЦКП в соответствии с требованиями, установленными в технической документации на изделие (поддержание требуемого температурно-влажностного режима, обеспечение требуемых параметров качества сети электропитания, сохранение заданных параметров заземления и др.).

Анализ задач технического обслуживания и ремонта позволяет выделить следующие направления работ по обеспечению надежности функционирования аппаратно-программных средств ЦКП.

1. Обеспечение работоспособности технических средств. При решении данной задачи необходимо осуществлять их диагностику, анализ и прогнозирование состояния.
2. Обеспечение работоспособности операционных систем и прикладного программного обеспечения. Проведение тестирования программных средств должно осуществляться таким образом, чтобы не только обнаруживать, но и предупреждать дефекты в их составе.
3. Обеспечение целостности, сохранности и работоспособности информационных массивов. Данная задача сводится к резервному архивированию данных и обеспечению их защиты от вирусов.
4. Обеспечение работоспособности периферийного, сетевого и коммуникационного оборудования.
5. Сбор, анализ и обработка статистических данных об отказах, сбоях и других аварийных ситуациях, связанных с техническими и программными средствами.

Методы обеспечения работоспособности аппаратно-программных средств можно разделить на активные и пассивные.

К основным *активным* методам обеспечения работоспособности АПС относятся:

- *удаление пыли и загрязнений* с вентиляционных устройств технических средств и компонентов их внутреннего монтажа;
- *резервное копирование системы*, позволяющее восстановить работоспособность при крупном аппаратно-программном сбое (отказе);
- *диагностика и обслуживание накопителей на жестких магнитных дисках* (НЖМД) на основе тестирования их физических поверхностей с применением известных утилит, по результатам тестирования НЖМД прогнозирование возможности продолжения их эксплуатации до очередного технического обслуживания;
- *дефрагментация файлов*;
- *своевременное обновление баз* в составе антивирусных средств.

Пассивные методы обычно направлены на защиту АПС от внешних неблагоприятных воздействий. Речь идет об установке защитных устройств в сети электропитания, поддержании чистоты и приемлемых температурно-влажностных характеристик в помещениях, где установлены АПС, уменьшении уровня вибрации, сохранении заданных параметров заземления и т. п.

4 Заключение

1. Показатели надежности ГВВК и систем, обеспечивающих его функционирование (электропитания и кондиционирования), составляют величины одного порядка, что отвечает требованиям по обеспечению равнопрочности элементов ЦКП.
2. Обеспечение надежности функционирования ЦКП ФИЦ ИУ РАН на достигнутом уровне в процессе эксплуатации возможно за счет:
 - выполнения установленных в технических описаниях и инструкциях по эксплуатации условий и правил применения технических и программных средств;
 - своевременного и качественного проведения технического обслуживания, включая диагностику технических и программных средств в целях раннего выявления причин отказов и принятия мер по их устранению;
 - резервного копирования системы;
 - тщательного проведения сбора, анализа и обработки статистических данных об отказах, сбоях и других аварийных ситуациях, связанных

с техническими и программными средствами с целью выявления причин наиболее частых неисправностей, прогнозирования состояния АПС, программного обеспечения и выдачи обоснованных рекомендаций по возможности продолжения их эксплуатации на период до очередного технического обслуживания.

Литература

1. *Зацаринный А. А., Гаранин А. И., Кондрашев В. А., Волович К. И., Мальковский С. И.* Оценка надежности гибридного высокопроизводительного вычислительного комплекса при решении научных задач // Системы и средства информатики, 2019. Т. 29. № 2. С. 135–147.
2. *Зацаринный А. А., Гаранин А. И., Козлов С. В.* Научно-практические аспекты обеспечения надежности информационно-телекоммуникационных сетей. — М.: ФИЦ ИУ РАН, 2017. 246 с.
3. *Беляев Ю. К., Богатырев В. А., Болотин В. В. и др.* Надежность технических систем: Справочник / Под ред. И. А. Ушакова. — М.: Радио и связь, 1985. 608 с.
4. *Кашианов В. А., Медведев А. И.* Теория надежности сложных систем. — М.: Физматлит, 2010. 608 с.

Поступила в редакцию 22.12.20

ENSURING THE RELIABILITY OF THE FRC CSC RAS CENTER FOR COLLECTIVE USE

A. A. Zatsarinny, A. I. Garanin, and S. A. Denisov

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

Abstract: The reliability of the center for collective use of the FRC CSC RAS is evaluated, which takes into account the reliability indicators of the following system’s component: a hybrid high-performance complex, power supply, and air conditioning systems. The analysis of the ways to improve the reliability of the center is carried out. The main conditions for ensuring the reliability of functioning and maintaining the operability of the hardware and hardware-software tools of the collective use center during their operation are formulated. Recommendations are given to ensure high-quality maintenance and repair of equipment.

Keywords: collective use center; power supply system; air conditioning system; reliability; failure; equivalent circuit for reliability calculation

DOI: 10.14357/08696527210203

Acknowledgments

The paper was partially supported by the Russian Foundation for Basic Research (projects 18-29-03091 and 18-29-03100).

References

1. Zatsarinny, A. A., A. I. Garanin, V. A. Kondrashev, K. I. Volovich, and S. I. Mal'kovsky. 2019. Otsenka nadezhnosti gibridnogo vysokoproizvoditel'nogo vychislitel'nogo kompleksa pri reshenii nauchnykh zadach [Evaluation of the reliability of a hybrid high-performance computing complex in solving scientific problems]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(2):135–147.
2. Zatsarinny, A. A., A. I. Garanin, and S. V. Kozlov. 2017. *Nauchno-prakticheskie aspekty obespecheniya nadezhnosti informatsionno-telekommunikatsionnykh setey* [Scientific and practical aspects of ensuring the reliability of information and telecommunication networks]. Moscow: FRC CSC RAS. 246 p.
3. Belyaev, Yu. K., V. A. Bogatyrev, V. V. Bolotin, *et al.* 1985. *Nadezhnost' tekhnicheskikh sistem* [Reliability of technical systems]. Moscow: Radio i svyaz'. 608 p.
4. Kashtanov, V. A., and A. I. Medvedev. 2010. *Teoriya nadezhnosti slozhnykh sistem* [Theory of complex systems reliability]. Moscow: Fizmatlit. 608 p.

Received December 22, 2020

Contributors

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

Garanin Alexander I. (b. 1951) — 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; Algaranin@mail.ru

Denisov Sergey A. (b. 1985) — leading engineer, 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; SDenisov@frcsc.ru

МЕТОДЫ АНАЛИЗА ДАННЫХ ЭЛЕКТРОЭНЦЕФАЛОГРАФИИ С ПРИМЕНЕНИЕМ СВЕРТОЧНЫХ И РЕКУРРЕНТНЫХ НЕЙРОННЫХ СЕТЕЙ*

И. А. Шанин¹, С. А. Ступников²

Аннотация: Современные методы анализа нейрофизиологических данных позволяют решать разнообразные задачи как в медицинской отрасли, так и в области построения интерфейсов «мозг–компьютер». В рамках данной работы были рассмотрены особенности решения задач анализа данных электроэнцефалографии (ЭЭГ) на примере задач детектирования артефактов и распознавания человеческих эмоций. Благодаря развитию алгоритмов на основе искусственных нейронных сетей, а также удешевлению промышленных интерфейсов эффективность и робастность современных методов анализа данных ЭЭГ приближается к уровню, достаточному для использования вне лабораторных условий. В работе предложены методы анализа данных ЭЭГ на основе сверточных и рекуррентных нейронных сетей, позволяющие добиться достаточно высокой точности классификации артефактов и распознавания эмоций на открытых наборах данных.

Ключевые слова: нейрофизиология; нейроинформатика; электроэнцефалография; анализ данных; нейронные сети; обнаружение артефактов в данных; детектирование эмоций

DOI: 10.14357/08696527210204

1 Введение

Электроэнцефалография — один из наиболее популярных инструментов регистрации активности головного мозга человека. Данные эксперимента электроэнцефалографии, как правило, представлены в виде многомерного временного ряда с частотой дискретизации от 100 Гц до 1 кГц, значения которого соответствуют электрическим потенциалам, измеренным на поверхности головы испытуемого. Медицинский анализ данных ЭЭГ головного мозга используется при диагностике эпилепсии, деменции, болезни Альцгеймера, прочих поражений головного мозга.

По сравнению с другими популярными методами, такими как функциональная магнитно-резонансная томография и позитронно-эмиссионная томография,

* Работа выполнена при финансовой поддержке РФФИ (проект 18-29-22096).

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

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

ЭЭГ обеспечивает высокое временное разрешение сигнала при более низком пространственном разрешении.

Отдельно выделяют потребительский класс устройств регистрации ЭЭГ-сигналов, оснащенных меньшим числом отведений и регистрирующих более шумный сигнал. Доступность таких устройств позволяет применять их вне лабораторных условий для создания интерфейса «мозг–компьютер». Таким образом, все меньшая часть доступных наборов данных поступает в рамках дорогостоящих медицинских экспериментов, в то время как все большую часть составляют данные, зарегистрированные коммерческими интерфейсами «мозг–компьютер».

Цель данной работы — провести исследование методов, применяемых для анализа данных ЭЭГ в нейрофизиологии, в том числе статистических методов и методов машинного обучения, в частности нейросетей глубокого обучения, на примере конкретных задач. Выбранные задачи детектирования артефактов в данных и распознавания эмоционального состояния человека особенно важны для создания аффективных систем взаимодействия человека и компьютера и систем эмоционального искусственного интеллекта.

Работа состоит из следующих разделов: разд. 2 посвящен постановкам задач анализа данных ЭЭГ; в разд. 3 рассматриваются существующие методы решения данных задач, включая методы извлечения признаков и методы на основе глубоких нейронных сетей; разд. 4 и 5 посвящены исследованию и разработке новых методов решения поставленных задач.

2 Задачи анализа данных на основе электроэнцефалографии

В обзорной работе [1] приведены задачи распознавания, наиболее успешно решаемые при использовании искусственных нейронных сетей: распознавание эмоций; распознавание моторной деятельности; оценка умственной нагрузки; детектирование эпилептических приступов; анализ фаз сна; анализ выявленных потенциалов. Сравнительная доступность технологий интерфейса «мозг–компьютер» делает возможным решение разного рода биометрических задач, таких как распознавание пола и возраста субъекта [2] и верификация личности [3]. Особенность таких задач заключается в необходимости создания наборов данных, содержащих ЭЭГ-сигналы существенно большего числа субъектов, чем обычно доступно в наборах данных, собранных в лабораторных условиях. Ниже более подробно рассмотрены две задачи анализа данных ЭЭГ.

2.1 Детектирование и устранение артефактов в данных электроэнцефалографии

В связи с тем что ЭЭГ регистрирует электрические потенциалы различной природы, перед решением любых задач распознавания необходимо очистить данные от «артефактов»: компонентов сигнала, не связанных с мозговой деятельностью субъекта. Артефакты можно разделить на *физиологические*, которые

появляются вследствие физиологических процессов организма, таких как мышечная активность, сердечная деятельность, движения глаз, а также на *физические*, возникающие в результате технических погрешностей, таких как сетевая наводка и погрешности измерений при загрязненных контактах. Таким образом, задача включает детектирование наличия артефакта в конкретный момент времени и построение очищенного сигнала (устранение артефактов).

2.2 Распознавание эмоционального состояния человека

Распознавание человеческих эмоций — важная задача построения аффективных систем. Информация об эмоциональном состоянии субъекта может быть использована в медицинских и в нейромаркетинговых [4] приложениях. В рамках большинства исследований эмоциональные состояния характеризуются либо множеством категорий, либо условными численными показателями интенсивности эмоции (от слабой до сильной) и ее характера (от негативного до позитивного). Таким образом, задача распознавания эмоций на основе данных ЭЭГ может рассматриваться как задача классификации или регрессии в зависимости от выбора способа разметки в конкретном эксперименте.

3 Известные методы решения задач анализа данных электроэнцефалографии

Современные методы решения задач анализа данных ЭЭГ на основе методов машинного обучения включают следующие этапы: предобработка данных; извлечение признаков; построение глубоких представлений; применение классификатора или регрессора. Параметры таких систем, как правило, определяются весами искусственных нейронных сетей, которые настраиваются с помощью решения единой задачи оптимизации функции потерь.

В работе [5] приведен обзор информативных признаков, применимых в задаче распознавания эмоций. В число рассмотренных признаков вошли такие признаки, извлекаемые из временной области сигнала, как мощность, среднее, стандартное отклонение, дифференциальные признаки, а также признаки, основанные на подсчете смен знака отфильтрованного сигнала (*higher order crossings*, НОС). В рамках анализа сигнала в частотной области рассматривается мощность сигнала в заданных частотных диапазонах, называемых ритмами, общепринятыми в задачах анализа данных ЭЭГ: альфа-ритм (8–12 Гц), бета-ритм (12–30 Гц), гамма-ритм (30–64 Гц), дельта-ритм (1–4 Гц) и тета-ритм (4–8 Гц), а также признаки, основанные на спектре высшего порядка. Особый интерес представляют признаки, выделяемые из временно-частотной области с использованием дискретного вейвлетного преобразования, преобразования Гильберта–Хуанга и оконное преобразование Фурье [5, разд. 2.2–2.3]. В рамках сравнительного анализа наиболее эффективными показали себя методы, основанные на НОС, спектре высшего порядка и на преобразовании Гильберта–Хуанга. Помимо признаков,

выделяемых независимо из каждого канала, также рассматривают топологическую структуру каналов электроэнцефалограмм, исследуя взаимосвязи между компонентами сигнала. Взаимосвязи между сигналами моделируются в виде графа, вершинами которого служат сигналы, а веса ребер соответствуют мере близости этих сигналов. Подходы, основанные на моделировании взаимосвязей между каналами ЭЭГ, применяются как для устранения артефактов [6], так и для распознавания эмоций [7]. В приведенных подходах извлечение информации о взаимосвязи между каналами происходит посредством применения обучаемого механизма внимания на графах, являющегося составным элементом искусственной нейронной сети.

3.1 Методы детектирования и устранения артефактов

В обзоре [8] рассмотрены основные подходы к задаче предварительной обработки и очистки сигнала ЭЭГ. К наиболее популярным методам относятся методы регрессионного анализа (для их применения дополнительно требуются данные электроокулографии или электрокардиографии), методы слепого разделения источников, в частности метод независимых компонент в комбинации с дискретным вейвлет-преобразованием, канонический корреляционный анализ и методы цифровой обработки сигналов. Рассмотрение задачи анализа ЭЭГ-сигнала как слепого разделения источников позволяет также применять глубокие нейронные сети на основе автокодировщиков. В работе [9] метод независимых компонент был заменен частью архитектуры нейронной сети, состоящей из автокодировщиков. Такой подход позволил в рамках одной нейросетевой архитектуры решить и задачу устранения артефактов, и задачу распознавания эмоций.

3.2 Методы распознавания эмоционального состояния по данным электроэнцефалографии

В последние годы задача распознавания эмоций стала популярным предметом исследования на пересечении нейрофизиологии и алгоритмов машинного обучения. Во многом это произошло благодаря появлению качественных наборов данных, таких как DEAP [10] и SEED [11], используемых как для обучения, так и для валидации моделей машинного обучения. В обзоре [12] перечислены наиболее популярные подходы на основе искусственных нейронных сетей, применяемых в рассматриваемой задаче. Среди методов, позволяющих достичь оптимального качества, применяются такие архитектуры нейросетей, как иерархическая сверточная нейронная сеть (CNN — convolutional neural network), вариант сверточной сети SincNet. Также хорошо себя зарекомендовали рекуррентные нейронные сети, в частности модели долгой краткосрочной памяти (LSTM — long short-term memory). Данные подходы позволяют добиться на наборах данных SEED и DEAP точности распознавания в диапазоне от 85% до 95%.

4 Детектирование артефактов электроэнцефалографии на основе комбинаций сверточных и рекуррентных нейронных сетей с использованием информации о взаимодействии между каналами

В этом разделе рассмотрен метод, нацеленный на эффективное детектирование основных типов артефактов сигналов ЭЭГ, таких как движения глаз, жевание, вздрагивание, дефекты контактных площадок и мышечные артефакты.

Основные этапы метода приведены на рис. 1. На этапе *извлечения признаков* происходит отбор каналов, представленных для достаточного числа субъектов, приведение сигналов к единой частоте дискретизации, спектральная фильтрация, разделение сигнала на пересекающиеся окна равной длины. В полученных окнах ЭЭГ-сигнала содержатся каналы, оставшиеся после отбора. Из этих каналов строится матрица попарных корреляций с помощью подсчета коэффициента ранговой корреляции Спирмена. Согласно гипотезе, попарная корреляция каналов ЭЭГ при наличии артефактов является важным информативным признаком, характеризующим природу наблюдаемого сигнала. Посчитанные таким образом матрицы служат входными данными искусственной нейронной сети, состоящей из *двумерного сверточного слоя* и операции снижения размерности средствами *пулингового слоя* (подробное описание основных элементов архитектур глубоких нейронных сетей, применяемых в анализе нейрофизиологических данных, при-

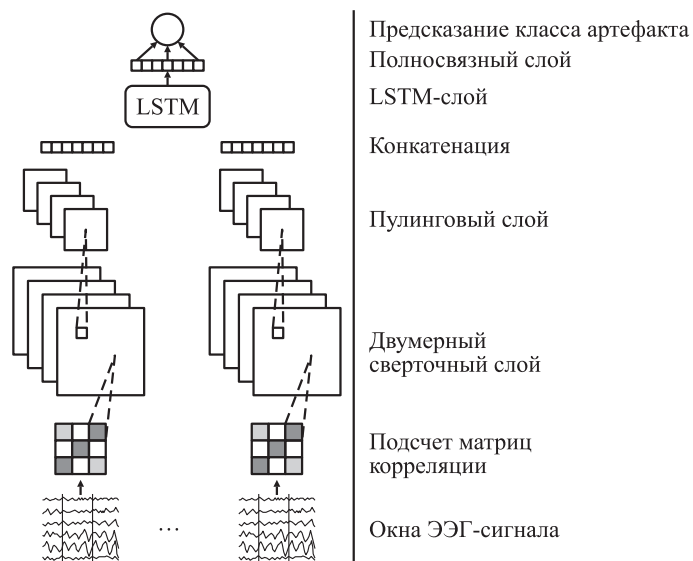


Рис. 1 Детектирование артефактов в данных ЭЭГ на основе матриц корреляций, архитектура искусственной нейронной сети

Детектирование артефактов на корпусе данных TUN

Вид модели	F1-мера	Точность
CNN-LSTM, фильтр, 16 окон	0,67	0,71
CNN-LSTM, без фильтра, 4 окна	0,56	0,73
CNN-LSTM, фильтр, 8 окон	0,7	0,71
CNN, фильтр, 8 окон	0,64	0,66
CNN, без фильтра, 8 окон,	0,67	0,69
XGBoost, 8 окон	0,67	0,69

ведено в работе [12]). Преобразованные таким образом данные представлены в виде вектора фиксированной размерности. Векторы нескольких последовательных окон используются в качестве входных данных *слоя долгой краткосрочной памяти* (LSTM) с последующим применением *полносвязного слоя*. Полученный вектор интерпретируется как вероятность отнесения последовательности из рассматриваемых окон к одному из исследуемых классов артефактов. Все слои искусственной нейронной сети обучаются совместно с использованием метода оптимизации Adam. В отличие от работы [6] сверточный слой применяется к матрице попарных корреляций без использования механизма внимания, а потому полученная модель менее склонна к переобучению, которое часто становится серьезной проблемой на наборах данных небольшого объема.

Метод был программно реализован¹ и опробован на наборе данных TUN EEG Artifact Corpus (TUAR) [13]. На этапе обработки использовался полосовой частотный фильтр с диапазоном 4–45 Гц. После очистки данных для анализа были отобраны 22 канала. Были рассмотрены 3-секундные окна с перекрытием 25%. Анализ проводился по 8 или по 16 последовательным окнам. Также опробован метод, исключая LSTM-слой.

По результатам тестирования (см. таблицу) лучшим оказался вариант, сочетающий сверточные и рекуррентные слои, использующий полосовой фильтр на этапе предобработки и использующий для анализа 8 последовательных окон. Данный метод показал себя лучше по F1-мере, что в условиях дисбаланса классов считается более приоритетной метрикой, нежели точность. Методы на основе нейросетей показали себя лучше, чем решение на основе ансамблей решающих деревьев XGBoost [14] (используемого в качестве базового метода для сравнения) как по точности, так и по F1-мере.

5 Робастное распознавание эмоций с применением многозадачного обучения

Одним из препятствий к созданию универсального метода распознавания эмоций по ЭЭГ является высокая межсубъектная вариативность, связанная, как

¹<https://github.com/shanin/eeg-dl-artifact-classification>.

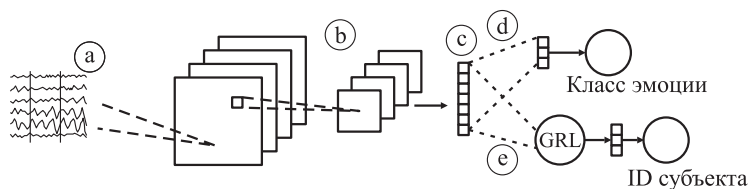


Рис. 2 Глубокая связательная нейронная сеть для задачи робастного распознавания эмоций на основе ЭЭГ

правило, с различающимися условиями регистрации ЭЭГ-сигналов и с индивидуальными особенностями работы головного мозга у субъектов. Таким образом, для построения робастного (т. е. устойчивого к изменению особенностей записи, различиям в характеристиках субъектов) метода распознавания требуется строить представления ЭЭГ-сигналов, не зависящие от межсубъектной вариативности данных. Отдельной проблемой является недостаточный объем размеченных данных: в наиболее популярных наборах данных содержится информация лишь о нескольких десятках субъектов.

Для преодоления данных проблем предлагается метод на основе состязательного многозадачного обучения, в рамках которого используется не только задача распознавания эмоций, но и задача классификации субъекта. Состязательный подход [15] заключается в том, что при обучении нейронной сети используется дополнительный классификатор, соответствующая ему функция потерь и операция инверсии градиента. В результате оптимизации функций потерь полученное решение стремится к наибольшей точности распознавания эмоций при одновременно наименьшей точности классификации субъектов, тем самым обеспечивая устойчивость решения к изменению домена.

Предлагаемый метод состоит из последовательного применения следующих операций (рис. 2): *извлечение спектральных признаков* с помощью оконного быстрого преобразования Фурье (a), *последовательное применение сверточных слоев* к полученной спектрограмме (b), *усреднение полученных глубинных представлений по временному измерению* (c), при этом размерность полученных представлений не зависит от длины входного сигнала. Затем полученные представления используются как входные значения для двух классификаторов, представленных в виде полносвязных слоев нейронной сети: *классификатора эмоций* (d) и *классификатора субъектов* (e). При этом классификатор субъектов предваряет операция инверсии градиента (Gradient Reversal Layer) [15]. В качестве функции потерь используется взвешенная сумма перекрестной энтропии классификации эмоциональных состояний и перекрестной энтропии классификации субъектов.

С помощью предложенного метода удалось добиться точности 79% на наборе данных DEAP [10], что сопоставимо с результатами, представленными в работе [12]. Ранее проблема доменной адаптации в задаче распознавания эмоций по ЭЭГ была рассмотрена в работе [16]. Основное отличие предложенного

метода заключается в том, что в рамках многозадачного обучения предусмотрена возможность использовать наборы данных, не содержащие информацию об эмоциональных состояниях. Таким образом, при обучении искусственной нейронной сети часть данных используется только для обучения классификатора субъекта. Разработка метода обучения такой модели входит в планы дальнейших исследований авторов настоящей работы.

6 Заключение

В работе проведено исследование методов анализа данных ЭЭГ на примере конкретных задач. На основании обзора родственных работ были предложены методы решения задачи детектирования артефактов в данных ЭЭГ и распознавания эмоционального состояния на основе искусственных нейронных сетей. Предложенные методы были протестированы на известных наборах данных, были получены результаты, сопоставимые с лучшими известными в литературе результатами для рассмотренных задач.

Литература

1. *Craik A., He Y., Contreras-Vidal J. L.* Deep learning for electroencephalogram (EEG) classification tasks: A review // *J. Neural Eng.*, 2019. Vol. 16. Iss. 3. Art. ID: 031001. 28 p. doi: 10.1088/1741-2552/ab0ab5.
2. *Kaur B., Singh D., Roy P. P.* Age and gender classification using brain-computer interface // *Neural Comput. Appl.*, 2019. Vol. 31. Iss. 10. P. 5887–5900. doi: 10.1007/s00521-018-3397-1.
3. *Gui Q., Ruiz-Blondet M. V., Laszlo S., Jin Z.* A survey on brain biometrics // *ACM Comput. Surv.*, 2019. Vol. 51. Iss. 6. P. 1–38. doi: 10.1145/3230632.
4. *Yadava M., Kumar P., Saini R., Roy P. P., Dogra D. P.* Analysis of EEG signals and its application to neuromarketing // *Multimed. Tools Appl.*, 2017. Vol. 76. Iss. 18. P. 19087–19111.
5. *Jenke R., Peer A., Buss M.* Feature extraction and selection for emotion recognition from EEG // *IEEE T. Affect. Comput.*, 2014. Vol. 5. Iss. 3. P. 327–339. doi: 10.1109/TAFFC.2014.2339834.
6. *Mauri G., Stella F., Morreale A., Cisotto G., Manzoni S., Zanga A., Zoppis I.* An attention-based architecture for EEG classification // *13th Joint Conference (International) on Biomedical Engineering Systems and Technologies Proceedings*, 2020. Vol. 4. P. 214–219.
7. *Song T., Zheng W., Song P., Cui Z.* EEG emotion recognition using dynamical graph convolutional neural networks // *IEEE T. Affect. Comput.*, 2018. Vol. 11. Iss. 3. P. 532–541. doi: 10.1109/TAFFC.2018.2817622.
8. *Jiang X., Bian G. B., Tian Z.* Removal of artifacts from EEG signals: A review // *Sensors*, 2019. Vol. 19. Iss. 5. Art. No. 987. 18 p. doi: 10.3390/s19050987.
9. *Xing X., Li Z., Xu T., Shu L., Hu B., Xu X.* SAE+ LSTM: A new framework for emotion recognition from multi-channel EEG // *Front. Neurobotics*, 2019. Vol. 13. Art. No. 37. 14 p. doi: 10.3389/fnbot.2019.00037.

10. Koelstra S., Muhl C., Soleymani M., Lee J. S., Yazdani A., Ebrahimi T., Patras I. Deap: A database for emotion analysis; using physiological signals // IEEE T. Affect. Comput., 2011. Vol. 3. Iss. 1. P. 18–31. doi: 10.1109/T-AFFC.2011.15.
11. Duan R. N., Zhu J. Y., Lu B. L. Differential entropy feature for EEG-based emotion classification // 6th IEEE/EMBS Conference (International) on Neural Engineering Proceedings. — Piscataway, NJ, USA: IEEE, 2013. P. 81–84. doi: 10.1109/NER.2013.6695876.
12. Gong S., Xing K., Cichocki A., Li J. Deep learning in EEG: Advance of the last ten-year critical period // arXiv.org, 22 Nov 2020. arXiv:2011.11128 [eess.SP].
13. Harati A., Lopez S., Obeid I., Picone J., Jacobson M. P., Tobochnik S. The TUH EEG Corpus: A big data resource for automated EEG interpretation // IEEE Signal Processing in Medicine and Biology Symposium Proceedings. — Piscataway, NJ, USA: IEEE, 2014. Art. ID: 7002953. 5 p. doi: 10.1109/SPMB.2014.7002953.
14. Chen T., Guestrin C. Xgboost: A scalable tree boosting system // 22nd ACM SIGKDD Conference (International) on Knowledge Discovery and Data Mining Proceedings. — New York, NY, USA: ACM, 2016. P. 785–794.
15. Ganin Y., Ustinova E., Ajakan H., Germain P., Larochelle H., Laviolette F., Lempitsky V. Domain-adversarial training of neural networks // J. Mach. Learn. Res., 2016. Vol. 17. Iss. 1. P. 2096–2030.
16. Lan Z., Sourina O., Wang L., Scherer R., Müller-Putz G. R. Domain adaptation techniques for EEG-based emotion recognition: A comparative study on two public datasets // IEEE T. Cognitive Developmental Systems, 2018. Vol. 11. Iss. 1. P. 85–94. doi: 10.1109/TCDS.2018.2826840.

Поступила в редакцию 27.12.20

ELECTROENCEPHALOGRAPHY DATA ANALYSIS WITH CONVOLUTIONAL AND RECURRENT NEURAL NETWORKS

I. A. Shanin and S. A. Stupnikov

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

Abstract: Modern methods for the neurophysiological data analysis provide promising solutions to various problems both in the field of medical industry and in the field of brain–computer interfaces development. In this paper, a couple of important electroencephalography (EEG) data analysis problems are considered that are artifact detection and removal and human emotion recognition. Due to recent active development of algorithms based on deep artificial neural networks and to the cost reduction of commercial prototypes of brain–computer interfaces, the efficiency and robustness of modern methods for EEG data analysis is approaching a level sufficient for use outside the laboratory. The paper proposes methods for EEG data analysis based on convolutional and recurrent neural

networks which make it possible to achieve high accuracy of artifact classification and emotion recognition over open data sets.

Keywords: neurophysiology; neuroinformatics; electroencephalography; data analysis artificial neural networks; data artifact detection; emotion recognition

DOI: 10.14357/08696527210204

Acknowledgments

This research was partially supported by the Russian Foundation for Basic Research (project 18-29-22096).

References

1. Craik, A., Y. He, and J. L. Contreras-Vidal. 2019. Deep learning for electroencephalogram (EEG) classification tasks: A review. *J. Neural Eng.* 16(3):031001. 28 p. doi: 10.1088/1741-2552/ab0ab5.
2. Kaur, B., D. Singh, and P. P. Roy. 2019. Age and gender classification using brain-computer interface. *Neural Comput. Appl.* 31(10):5887–5900. doi: 10.1007/s00521-018-3397-1.
3. Gui, Q., M. V. Ruiz-Blondet, S. Laszlo, and Z. Jin. 2019. A survey on brain biometrics. *ACM Comput. Surv.* 51(6):1–38. doi: 10.1145/3230632.
4. Yadava, M., P. Kumar, R. Saini, P. P. Roy, and D. P. Dogra. 2017. Analysis of EEG signals and its application to neuromarketing. *Multimed. Tools Appl.* 76(18):19087–19111.
5. Jenke, R., A. Peer, and M. Buss. 2014. Feature extraction and selection for emotion recognition from EEG. *IEEE T. Affect. Comput.* 5(3):327–339. doi: 10.1109/TAFFC.2014.2339834.
6. Mauri, G., F. Stella, A. Morreale, G. Cisotto, S. Manzoni, A. Zanga, and I. Zoppis. 2020. An attention-based architecture for EEG classification. *13th Joint Conference (International) on Biomedical Engineering Systems and Technologies Proceedings.* 4:214–219.
7. Song, T., W. Zheng, P. Song, and Z. Cui. 2018. EEG emotion recognition using dynamical graph convolutional neural networks. *IEEE T. Affect. Comput.* 11(3):532–541. doi: 10.1109/TAFFC.2018.2817622.
8. Jiang, X., G. B. Bian, and Z. Tian. 2019. Removal of artifacts from EEG signals: A review. *Sensors* 19(5):987. 18 p. doi: 10.3390/s19050987.
9. Xing, X., Z. Li, T. Xu, L. Shu, B. Hu, and X. Xu. 2019. SAE+ LSTM: A new framework for emotion recognition from multi-channel EEG. *Front. Neurobotics* 13:37. 14 p. doi: 10.3389/fnbot.2019.00037.
10. Koelstra, S., C. Muhl, M. Soleymani, J. S. Lee, A. Yazdani, T. Ebrahimi, and I. Patras. 2011. Deap: A database for emotion analysis; using physiological signals. *IEEE T. Affect. Comput.* 3(1):18–31. doi: 10.1109/T-AFFC.2011.15.
11. Duan, R. N., J. Y. Zhu, and B. L. Lu. 2013. Differential entropy feature for EEG-based emotion classification. *6th IEEE/EMBS Conference (International) on Neural Engineering Proceedings.* Piscataway, NJ: IEEE. 81–84. doi: 10.1109/NER.2013.6695876.

12. Gong, S., K. Xing, A. Cichocki, and J. Li. 2020. Deep learning in EEG: Advance of the last ten-year critical period. arXiv:2011.11128 [eess.SP]. Available at: <https://arxiv.org/abs/2011.11128> (accessed April 5, 2021).
13. Harati, A., S. Lopez, I. Obeid, J. Picone, M. P. Jacobson, and S. Tobochnik. 2014. The TUH EEG corpus: A big data resource for automated EEG interpretation. *Signal Processing in Medicine and Biology Symposium Proceedings*. Piscataway, NJ: IEEE. Art. ID 7002953. 5 p. doi: 10.1109/SPMB.2014.7002953.
14. Chen, T., and C. Guestrin. 2016. Xgboost: A scalable tree boosting system. *22nd ACM SIGKDDI Conference (International) on Knowledge Discovery and Data Mining Proceedings*. New York, NY: ACM. 785–794.
15. Ganin, Y., E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, and V. Lempitsky. 2016. Domain-adversarial training of neural networks. *J. Mach. Learn. Res.* 17(1):2096–2030.
16. Lan, Z., O. Sourina, L. Wang, R. Scherer, and G. R. Müller-Putz. 2018. Domain adaptation techniques for EEG-based emotion recognition: A comparative study on two public datasets. *IEEE T. Cognitive Developmental Systems* 11(1):85–94. doi: 10.1109/TCDS.2018.2826840.

Received December 27, 2020

Contributors

Shanin Ivan A. (b. 1991) — junior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; ishanin@frccsc.ru

Stupnikov Sergey A. (b. 1978) — 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; [sstupnikov@ipiran.ru](mailto:ssupnikov@ipiran.ru)

СТИМУЛЯЦИЯ КОНФЛИКТОВ АГЕНТОВ В ГИБРИДНЫХ ИНТЕЛЛЕКТУАЛЬНЫХ МНОГОАГЕНТНЫХ СИСТЕМАХ

С. В. Листонад¹, И. А. Кириков²

Аннотация: Конфликты в коллективах специалистов, решающих проблемы «за круглым столом», — неотъемлемая часть процесса обсуждения, и при компьютерном моделировании этого процесса необходимо управлять конфликтами по аналогии с тем, как это происходит в оригинале. Не каждый конфликт носит деструктивный характер и требует подавления: управление конфликтом в коллективе предполагает идентификацию ситуации принятия решений, при необходимости стимуляцию и последующее разрешение конструктивных форм конфликта, а также предотвращение его деструктивных форм. Статья посвящена разработке метода стимуляции конфликтов между агентами в гибридных интеллектуальных многоагентных системах (ГиИМАС), что позволит разработать подсистему управления конфликтами, обеспечив релевантность таких систем малым коллективам специалистов, успешно решающим проблемы с высокой комбинаторной сложностью, неоднородностью, недоопределенностью и другими НЕ-факторами.

Ключевые слова: конфликт; гибридная интеллектуальная многоагентная система; коллектив специалистов; стимулирование конфликта

DOI: 10.14357/08696527210205

1 Введение

В практике управления производственными, экономическими и социальными системами широкий класс проблем традиционно решается коллективами специалистов. Коллективное решение проблем предпочтительней перед индивидуальным, если проблема обладает полителичностью, может иметь несколько альтернативных решений, затрагивает интересы различных сторон и важно учесть потребности каждой из них [1, 2]. Обратная сторона всестороннего рассмотрения проблемы коллективом специалистов — возникновение между ними конфликтов различного характера и необходимость задействовать механизмы управления ими для обеспечения эффективной работы коллектива. Не всякие конфликты требуется предотвращать: без конфликтов невозможна самоорганизация коллективов, предполагающая самостоятельное, без принуждающего влияния извне, формирование их состава, структуры и свойств [3]. В этой связи актуальна

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

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

стратегия стимуляции конфликтов, направленная на мобилизацию активности членов коллектива, развитие позитивной инициативы [4].

Другой недостаток коллективного решения проблем — длительное время выработки решений. Если проблема характеризуется не только полнотой, но и высокой динамичностью, возникают временные ограничения на длительность процедуры ее решения, из-за которых привлечение коллектива специалистов нецелесообразно. Для таких проблем в [5] предложены гибридные интеллектуальные многоагентные системы, агенты которых моделируют рассуждения отдельных специалистов, а системы в целом отображают в памяти компьютера макроуровневые процессы, возникающие в результате их взаимодействия при решении проблемы «за круглым столом». Они интегрируют подход гибридных интеллектуальных систем А. В. Колесникова [6] и аппарат многоагентных систем в смысле В. Б. Тарасова [7]. В [8] для управления конфликтами, возникающими между агентами, по аналогии с тем, как это происходит в коллективах специалистов, предложена модель ГиИМАС с проблемно- и процессно-ориентированными конфликтами. Цель настоящей работы — разработка метода стимуляции конфликтов между агентами в рамках данной модели.

2 Стимуляция конфликтов в коллективах специалистов

Один из элементов эффективного управления конфликтом в коллективах специалистов, работающих над решением проблемы «за круглым столом», — стимулирование конфликтов, когда это необходимо [9]. Управление конфликтом — целенаправленное, обусловленное объективными законами воздействие на его динамику в интересах развития или разрушения социальной системы, к которой относится конфликт [10]. При управлении конфликтами в малых коллективах специалистов актуален его продуктивный аспект, направленный на развитие коллектива и состоящий в предотвращении деструктивных конфликтов (преимущественно конфликтов по поводу отношений, личным вопросам и проблемам), стимулировании и разрешении конструктивных (инструментальных конфликтов относительно проблемы или процесса ее решения).

Традиционно считается, что конфликт подрывает цели существования коллектива или свидетельствует о неэффективности его работы. Однако отсутствие конфликта может означать эффект группинга, когда члены коллектива, несогласные с мнением большинства, не высказывают свою точку зрения и принимаются решение, не устраивающее некоторых или даже большинство членов.

В [11–14] показано, что умеренные инструментальные конфликты полезны: они стимулируют дискуссии и дебаты, помогающие коллективам повысить производительность за счет лучшего понимания различных точек зрения и альтернативных решений. В особенности этот эффект проявляется в коллективах, решающих не рутинные, а творческие, нестандартные проблемы. При этом конфликты, в том числе инструментальные, могут снизить групповую лояльность, приверженность коллективу и удовлетворенность работой [13].

Как показано в [15], коллективы стремятся избегать конфликтов, а при необходимости исключить кого-то из своих членов. Обычно исключают наиболее конфликтных, даже если есть понимание, что спровоцированные ими конфликты привели к повышению качества принимаемых решений. Если напряженность конфликта в коллективе очень низкая, фасилитатор, т. е. член коллектива, отвечающий за организацию эффективной совместной работы, должен предпринять действия по стимулированию инструментальных конфликтов, контролируя соблюдение норм поведения, препятствующих возникновению деструктивных персональных конфликтов по поводу отношений.

На основе анализа работ в области групповой динамики и конфликтологии [15–18] выделены и обобщены следующие условия стимуляции конфликта:

- (1) возможно развитие инструментального, а не персонального конфликта, например, если члены коллектива не склонны в ходе дискуссии по проблеме переходить на личности и проявлять излишнюю эмоциональность;
- (2) низкая напряженность конфликтов в коллективе, свидетельствующая о низкой активности и эффективности совместной работы. Причины низкой напряженности конфликтов: большое число специалистов-конформистов, коллектив нацелен не на поиск качественных решений, а на поиск компромисса и сохранение согласия членов, сопротивление членов коллектива переменам;
- (3) позитивная взаимозависимость целей участников коллектива, когда каждый из них не может достичь своих целей самостоятельно, без помощи других членов. В случае независимости или отрицательной взаимозависимости целей участников коллектива, когда достижение цели одним участником делает невозможным достижение целей остальными («победитель получает все»), стимулирование конфликта снижает эффективность работы коллектива.

Выделяются следующие меры стимуляции конфликтов в коллективах [15, 17, 19, 20]: привлечение в коллектив специалистов с разнородными ценностями, опытом и стилем работы; повышение конкуренции специалистов путем поощрения эффективной работы; создание среды новаторского, дивергентного мышления; назначение «адвоката дьявола»; перераспределение ролей в коллективе и его реструктуризация. Рассмотрим реализацию этих механизмов в ГиИМАС.

3 Управление конфликтами в гибридных интеллектуальных многоагентных системах

Подробное описание модели ГиИМАС с проблемно- и процессно-ориентированными конфликтами представлено в [8]. Рассмотрим основные ее элементы, необходимые для описания процесса управления конфликтами. С учетом того что в ГиИМАС эмоции специалистов не моделируются агентами и конфликт не может превратиться из инструментального в персональный, при управлении

конфликтами агентов учитываются только последние два из трех описанных выше условий необходимости стимуляции конфликта.

Формально ГиИМАС в целом определяется следующим образом [8]:

$$\text{himas} = \langle \text{AG}^*, \text{env}, \text{INT}, \text{ORG}, \text{MLP} \rangle, \quad (1)$$

где $\text{AG}^* = \{ag_1, \dots, ag_n, ag^{\text{dm}}, ag^{\text{fc}}\}$ — множество агентов, включающее n агентов-специалистов (АС) $ag_i, i \in \mathbb{N}, 1 \leq i \leq n$, агента, принимающего решения (АПР), — ag^{dm} и агента-фасилитатора (АФ), управляющего взаимодействиями агентов при решении проблемы и конфликтами между ними, — ag^{fc} ; env — концептуальная модель внешней среды системы; INT^* — множество элементов структурирования взаимодействий агентов [8], содержащее среди прочего модель предметной области ont ; ORG — множество архитектур ГиИМАС; $\text{MLP} = \{ \text{cnfm}, \text{gdid} \}$ — множество концептуальных моделей макроуровневых процессов в ГиИМАС, содержащее модель cnfm процесса управления конфликтами агентов и модель gdid идентификации взаимозависимости целей агентов.

Агент $ag_{id} \in \text{AG}^*$ из формулы (1) описывается выражением:

$$ag_{id} = \langle id_{id}^{\text{ag}}, gl_{id}^{\text{ag}}, \text{ACT}_{id}^{\text{ag}} \rangle,$$

где id_{id}^{ag} — идентификатор агента; gl_{id}^{ag} — нечеткая цель агента, т. е. нечеткое множество с функцией принадлежности $\mu_{id}(pr_1, \dots, pr_{\text{NPRG}})$, заданной на подмножестве целевых концептов-свойств $\text{PR}^g = \{pr_1, \dots, pr_{\text{NPRG}}\}$ множества концептов-свойств $\text{PR}^g \subseteq \text{PR}$ модели предметной области ont ; $\text{ACT}_{id}^{\text{ag}}$ — множество действий агента.

Модель процесса управления конфликтами агентов описывается выражением:

$$\text{cnfm} = \langle \text{CNF}, \text{cnfcl}, \text{ACT}^{\text{afcm}}, \text{ACT}^{\text{agcs}}, \text{ACT}^{\text{agcr}} \rangle, \quad (2)$$

где CNF — матрица конфликтов между парами агентов; cnfcl — классификатор конфликтов агентов [8], формирующий матрицу CNF , $\text{ACT}^{\text{afcm}} = \{ \text{act}^{\text{cnfm}}, \text{act}^{\text{cnfi}}, \text{act}^{\text{cnfs}}, \text{act}^{\text{cnfr}} \}$ — множество функций АФ по управлению конфликтами АС, содержащее функцию «управление конфликтом» act^{cnfm} , обеспечивающую идентификацию act^{cnfi} конфликтов с помощью классификатора cnfcl и инициализацию функции стимуляции act^{cnfs} или разрешения act^{cnfr} конфликтов; ACT^{agcs} — множество действий АС, выполняемых при стимулировании противоречий АФ, $\text{ACT}^{\text{agcs}} \subseteq \bigcup_{ag_{id} \in \text{AG}^*} \text{ACT}_{id}^{\text{ag}}$; ACT^{agcr} — множество допустимых действий АС по разрешению противоречий, $\text{ACT}^{\text{agcr}} \subseteq \bigcup_{ag_{id} \in \text{AG}^*} \text{ACT}_{id}^{\text{ag}}$.

Конфликт между агентами, т. е. элемент матрицы из выражения (2), описывается следующей формулой:

$$\text{cnf}_{i_j \text{cnft}} = \langle ag_i, ag_j, \text{cnfin}, \text{cnft}, \text{ACT}_i^{\text{agcr}}, \text{ACT}_j^{\text{agcr}} \rangle,$$

где ag_i и ag_j — агенты-субъекты конфликта, $i, j \in \mathbb{N}$, $1 \leq i, j \leq n$, $i \neq j$; $snfn$ — напряженность конфликта в виде скалярной величины $snfn \in [0, 1]$, вычисляемая классификатором конфликтов $snfcl$ в соответствии с реализуемой им мерой напряженности конфликта [8]; $snft$ — символьная переменная «тип конфликта», определенная на множестве $CNFT = \{\text{«проблемно-ориентированный»}, \text{«процессно-ориентированный»}\}$; ACT_i^{agcr} и ACT_j^{agcr} — множество допустимых действий агентов ag_i и ag_j соответственно по разрешению противоречий, $ACT_i^{agcr} \subseteq ACT_i^{ag}$, $ACT_j^{agcr} \subseteq ACT_j^{ag}$, $ACT_i^{agcr}, ACT_j^{agcr} \subseteq ACT^{agcr}$.

Модель идентификации взаимозависимости целей агентов описывается выражением:

$$gdid = \langle \mathbf{GD}, gdi \rangle.$$

Здесь \mathbf{GD} — матрица показателей взаимозависимости целей пар агентов, $gd_{ij} \in [-1, 1]$:

$$gd_{ij} = \begin{cases} -1 & \text{— максимальная отрицательная взаимозависимость;} \\ 0 & \text{— независимость;} \\ 1 & \text{— максимальная положительная взаимозависимость целей} \\ & \text{пары агентов } ag_i \text{ и } ag_j; \end{cases}$$

gdi — показатель взаимозависимости целей агентов, формирующий матрицу \mathbf{GD} в случае одномерных целей, т. е. $|\mathbf{PR}^g| = 1$ в соответствии с выражением

$$gd_{ij} = gdi(g_i^{ag}, g_j^{ag}) = \frac{\int_{v_{\min}}^{v_{\max}} \mu_{i \cap j}(\mathbf{pr}) d(\mathbf{pr})}{\int_{v_{\min}}^{v_{\max}} \mu_i(\mathbf{pr}) d(\mathbf{pr})} + \frac{\int_{v_{\min}}^{v_{\max}} \mu_{i \cap j}(\mathbf{pr}) d(\mathbf{pr})}{\int_{v_{\min}}^{v_{\max}} \mu_j(\mathbf{pr}) d(\mathbf{pr})} - 1, \quad (3)$$

где v_{\min} и v_{\max} — минимальное и максимальное значения свойства \mathbf{pr} ; $\mu_{i \cap j}(\mathbf{pr})$ — пересечение функций принадлежности нечетких целей агентов ag_i и ag_j .

Таким образом, функция «управление конфликтом» act^{cnfm} АФ может быть представлена следующей последовательностью шагов:

- (1) инициализировать переменную «стадия»: $stg = \text{«начало»}$;
- (2) используя выражение (3), вычислить среднее арифметическое показателей взаимозависимости целей агентов по формуле

$$gd^{himas} = \sum_{i=1}^n \sum_{j=i+1}^n 2 gd_{ij} (n-2)! (n!)^{-1};$$

- (3) запустить функцию идентификации конфликтов act^{cnfi} , чтобы с помощью классификатора конфликтов cnfcl на основе решений, предложенных АС, сформировать матрицу конфликтов **CNF** между парами агентов [8], после чего вычислить общий показатель напряженности конфликта в ГиИМАС

$$\text{cnf}^{\text{himas}} = \sum_{i=1}^n \sum_{j=i+1}^n \sum_{\text{cnft} \in \text{CNFT}} 2\text{cnf}_{ij\text{cnft}} (n-2)! |\text{CNFT}|^{-1} (n!)^{-1};$$

- (4) если $\text{gd}^{\text{himas}} > 0$, $\text{cnf}^{\text{himas}} < \text{cnf}_h^{\text{himas}}$ и $\text{stg} \in \{\text{«начало»}, \text{«стимуляция»}\}$, где $\text{cnf}_h^{\text{himas}}$ — определяемое в ходе тестирования системы значение, по умолчанию равное 0,5, то установить $\text{stg} = \text{«стимуляция»}$ и выполнить функцию «стимуляция конфликтов» act^{cnfs} , иначе установить $\text{stg} = \text{«разрешение»}$ и выполнить функцию «разрешение конфликтов» act^{cnfr} ;
- (5) если $\text{stg} = \text{«конец»}$ после выполнения функций «стимуляция конфликтов» act^{cnfs} или «разрешение конфликтов» act^{cnfr} , то отправить АПР сообщение о необходимости завершения работы ГиИМАС и закончить работу;
- (6) ожидать поступления сообщений от АС или АПР;
- (7) если получено сообщение от АС, содержащее решение проблемы или ее части, то перейти к п. 3;
- (8) если получено сообщение от АПР о завершении работы, закончить работу, иначе сообщить агенту-отправителю об ошибке и перейти к п. 5.

Поскольку работа посвящена вопросам стимуляции конфликтов в ГиИМАС, рассмотрим функцию act^{cnfs} подробнее.

4 Методы стимуляции конфликтов агентов в гибридных интеллектуальных многоагентных системах

Для стимуляции конфликтов в ГиИМАС предлагается использовать подходы, аналогичные применяемым в коллективах специалистов: моделирование агентами специалистов различных направлений с различными целями, применение методов дивергентного коллективного мышления [21], применение системы «вознаграждений» для усиления конкуренции, механизм «адвоката дьявола», при котором цели АС поочередно инвертируются.

При каждом вызове функции «стимуляция конфликтов» act^{cnfs} выполняется выбор метода коллективных рассуждений АС из набора доступных АФ. Длительность периода, в течение которого остается активным выбранный метод коллективных рассуждений АС, задается пользователем системы, по умолчанию она составляет $10n$ вызовов функции act^{cnfs} , где n — число АС. Выбор методов организован таким образом, чтобы очередной метод способствовал повышению

напряженности конфликта в ГиИМАС по сравнению с предыдущим: мозговой шторм с наводящими вопросами, пул мозговой записи, оспаривание ограничений, конкуренция за вознаграждение, инвертирования целей АС. Алгоритмы, реализующие первые три метода, представлены в [21], поэтому здесь они не рассматриваются.

Метод «конкуренции за вознаграждение» представляется следующей последовательностью шагов:

- (1) АПР рассылает АС постановку проблемы и предполагаемый вариант ее решения, если он есть;
- (2) АС на своих моделях предметной области генерируют варианты решений;
- (3) найденные решения собираются АПР;
- (4) АПР рассылает найденные решения АС, которые описывают достоинства своего решения и недостатки чужих решений, цель — максимизация числа баллов, полученных от АПР;
- (5) АПР начисляет баллы АС в соответствии с качеством его решений, которое оценивается как среднее арифметическое оценок, выставленных АПР и агентами-специалистами.

Метод «инвертирования целей АС» описывается следующими шагами:

- (1) АПР рассылает АС постановку проблемы;
- (2) АФ случайным образом в качестве «адвоката дьявола» выбирает одного АС из числа ранее не выбиравшихся; если таковых не осталось, алгоритм завершает работу;
- (3) по заданию АФ выбранный АС временно меняет свою цель на противоположную, генерирует решения в соответствии с новой целью и пересылает их АПР и остальным АС;
- (4) АС на своих моделях предметной области пытаются улучшить решения АС с инвертированной целью;
- (5) найденные решения собираются АПР;
- (6) АС с инвертированной целью меняет свою цель на исходную;
- (7) переход к п. 2.

Если при очередном запуске функции «стимуляция конфликтов» act^{cnfs} длительность работы по методу «инвертирования целей АС» превышает заданное пользователем или установленное по умолчанию ограничение, устанавливается $stg = \text{«конец»}$ и функция завершает работу. Это означает, что, несмотря на применение методов стимулирования конфликта, напряженность конфликтов в ГиИМАС осталась низкой, что свидетельствует о наличии очевидного решения проблемы, с которым согласны все АС, и необходимости завершения работы.

Таким образом, с использованием перечисленных методов АФ при необходимости может повысить напряженность конфликтов в ГиИМАС и разнообразие

рассматриваемых в ходе решения проблемы альтернатив. Система адаптируется к решаемой проблеме, заставляя АС рассматривать неочевидные варианты решения проблемы, если разнообразие предлагаемых альтернатив недостаточно, либо переходить к этапу согласования позиций АС, если между ними возникает большое число конфликтов высокой интенсивности. С учетом того что АС реализуют различные технологии искусственного интеллекта, ГиИМАС динамически вырабатывает гибридный интеллектуальный метод, релевантный очередной решаемой проблеме. В результате предложенная модель ГиИМАС, управляющая проблемно- и процессно-ориентированными конфликтами, а также методы стимулирования конфликтов повысят релевантность ГиИМАС малым коллективам специалистов, решающим проблемы «за круглым столом».

5 Заключение

Рассмотрены особенности, выделены условия и меры стимулирования конфликтов в коллективах специалистов, решающих практические проблемы «за круглым столом». Представлена модель ГиИМАС с проблемно- и процессно-ориентированными конфликтами, позволяющая имитировать процесс коллективного решения проблем на основе анализа напряженности конфликта, чтобы вырабатывать решения, сопоставимые с рекомендациями коллективов. Предложены алгоритм управления конфликтами в ГиИМАС и методы стимулирования конфликтов, аналогичные используемым в коллективах специалистов. Стимулирование конструктивных проблемно- и процессно-ориентированных конфликтов в ГиИМАС позволяет расширить множество рассматриваемых на последующих этапах решения проблемы альтернатив, учесть различные точки зрения и интересы специалистов, моделируемых агентами системы, повышая релевантность таких систем реальным коллективам специалистов.

Литература

1. Самсонова М. В., Ефимов В. В. Технология и методы коллективного решения проблем. — Ульяновск: УлГТУ, 2003. 152 с.
2. Спиридонов В. Ф. Психология мышления: решение задач и проблем. — М.: Генезис, 2006. 319 с.
3. Новосельцев В. И. Системный анализ: современные концепции. — Воронеж: Квартал, 2003. 360 с.
4. Хохлов А. С. Конфликтология. История. Теория. Практика. — Самара: СФ МГПУ, 2014. 312 с.
5. Колесников А. В., Кириков И. А., Листопад С. В. Гибридные интеллектуальные системы с самоорганизацией: координация, согласованность, спор. — М.: ИПИ РАН, 2014. 189 с.
6. Колесников А. В. Гибридные интеллектуальные системы. Теория и технология разработки. — СПб.: СПбГТУ, 2001. 711 с.

7. *Тарасов В. Б.* От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. — М.: Эдиториал УРСС, 2002. 352 с.
8. *Листонад С. В., Кириков И. А.* Метод идентификации конфликтов агентов в гибридных интеллектуальных многоагентных системах // Системы и средства информатики, 2020. Т. 30. № 1. С. 56–65. doi: 10.14357/08696527200105.
9. *Bauer T., Erdogan B.* Organizational behavior. — Irvington, NY, USA: Flat World Knowledge, 2009. 395 p.
10. *Емельянов С. М.* Практикум по конфликтологии. — СПб.: Питер, 2009. 384 с.
11. *Burgeois L. J.* Strategic goals, environmental uncertainty, and economic performance in volatile environments // Acad. Manage. J., 1985. Vol. 28. P. 548–573.
12. *Eisenhardt K., Schoonhoven C.* Organizational growth: Linking founding team, strategy, environment, and growth among U. S. semiconductor ventures, 1878–1988 // Admin. Sci. Quart., 1990. Vol. 35. P. 504–529.
13. *Jehn K. A., Northcraft G. B., Neale M. A.* Why differences make a difference: A field study of diversity, conflict, and performance in workgroups // Admin. Sci. Quart., 1999. Vol. 44. P. 741–763.
14. *Rahim M. A.* Toward a theory of managing organizational conflict // Int. J. Confl. Manage., 2002. Vol. 13. Iss. 3. P. 206–235.
15. *Sahu R. K.* Group dynamics & team building. — New Delhi: Excel Books, 2010. 632 p.
16. *Robbins S. P.* “Conflict management” and “conflict resolution” are not synonymous terms // Calif. Manage. Rev., 1978. Vol. 21. Iss. 2. P. 67–75.
17. *Van de Vliert E., de Dreu C. K. W.* Optimizing performance by conflict stimulation // Int. J. Confl. Manage., 1994. Vol. 5. Iss. 3. P. 211–222.
18. *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.
19. Management of conflict resolution & stimulation techniques. <https://theintactone.com/2019/03/14/ncm-u1-topic-7-management-of-conflict-resolution-stimulation-techniques>.
20. *Tanuja A.* Conflict: Features, causes and consequences | Organisation. <https://www.businessmanagementideas.com/notes/management-notes/conflict-management/conflict-features-causes-and-consequences-organisation/5009>.
21. *Колесников А. В., Листонад С. В.* Модель гибридной интеллектуальной многоагентной системы гетерогенного мышления для информационной подготовки оперативных решений в региональных электрических сетях // Системы и средства информатики, 2018. Т. 28. № 4. С. 31–41. doi: 10.14357/08696527180404.

Поступила в редакцию 02.03.21

STIMULATION OF AGENT CONFLICTS IN HYBRID INTELLIGENT MULTIAGENT SYSTEMS

S. V. Listopad and I. A. Kirikov

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

Abstract: Conflicts in teams of specialists solving the problems at a round table are an integral part of the discussion process and in computer modeling of this process, it is necessary to manage conflicts by analogy with how it happens in the original. Not every conflict is destructive and requires suppression: managing a conflict in a team involves identifying a decision-making situation, if necessary, stimulating and then resolving constructive forms of conflict, as well as preventing its destructive forms. The article is devoted to the development of a method for stimulating conflicts between agents in hybrid intelligent multiagent systems which will allow the development of a conflict management subsystem, ensuring the relevance of such systems to small teams of specialists who successfully solve the problems with high combinatorial complexity, heterogeneity, underdetermination, and other NON-factors.

Keywords: conflict; hybrid intelligent multiagent system; team of specialists; stimulating conflict

DOI: 10.14357/08696527210205

References

1. 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.
2. Spiridonov, V. F. 2006. *Psikhologiya myshleniya: Resheniye zadach i problem* [Psychology of thinking: Solving tasks and problems]. Moscow: Genezis. 319 p.
3. Novosel'tsev, V. I. 2003. *Sistemnyy analiz: Sovremennye kontseptsii* [System analysis: Modern concepts]. Voronezh: Kvarta. 360 p.
4. Khokhlov, A. S. 2014. *Konfliktologiya. Istoriya. Teoriya. Praktika* [Conflictology. History. Theory. Practice]. Samara: SF MGPU. 312 p.
5. Kolesnikov, A. V., I. A. Kirikov, and S. V. Listopad. 2014. *Gibridnye intellektual'nye sistemy s samoorganizatsiyey: koordinatsiya, soglasovannost', spor* [Hybrid intelligent systems with self-organization: Coordination, consistency, and dispute]. Moscow: IPI RAN. 189 p.
6. Kolesnikov, A. V. 2001. *Gibridnye intellektual'nye sistemy. Teoriya i tekhnologiya razrabotki* [Hybrid intelligent systems: Theory and technology of development]. St. Petersburg: SPbSTU Publs. 711 p.

7. Tarasov, V. B. 2002. *Ot mnogoagentnykh sistem k intellektualnym organizatsiyam: filozofiya, psikhologiya, informatika* [From multiagent systems to intelligent organizations: Philosophy, psychology, and informatics]. Moscow: Editorial URSS. 352 p.
8. Listopad, S. V., and I. A. Kirikov. 2020. Metod identifikatsii konfliktov agentov v gibrnidnykh intellektual'nykh mnogoagentnykh sistemakh [Agent conflict identification method in hybrid intelligent multiagent systems]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(1):56–65. doi: 10.14357/08696527200105.
9. Bauer, T., and B. Erdogan. 2009. *Organizational behavior*. Irvington, NY: Flat World Knowledge. 395 p.
10. Emel'yanov, S. M. 2009. *Praktikum po konfliktologii* [Tutorial at conflictology]. St. Petersburg: Piter. 384 p.
11. Burgeois, L. J. 1985. Strategic goals, environmental uncertainty, and economic performance in volatile environments. *Acad. Manage. J.* 28:548–573.
12. Eisenhardt, K., and C. Schoonhoven. 1990. Organizational growth: Linking founding team, strategy, environment, and growth among U.S. semiconductor ventures, 1878–1988. *Admin. Sci. Quart.* 35:504–529.
13. Jehn, K. A., G. B. Northcraft, and M. A. Neale. 1999. Why differences make a difference: A field study of diversity, conflict, and performance in workgroups. *Admin. Sci. Quart.* 44:741–763.
14. Rahim, M. A. 2002. Toward a theory of managing organizational conflict. *Int. J. Confl. Manage.* 13(3):206–235.
15. Sahu, R. K. 2010. *Group dynamics & team building*. New Delhi: Excel Books. 632 p.
16. Robbins, S. P. 1978. “Conflict management” and “conflict resolution” are not synonymous terms. *Calif. Manage. Rev.* 21(2):67–75.
17. Van de Vliert, E., and C. K. W. de Dreu. 1994. Optimizing performance by conflict stimulation. *Int. J. Confl. Manage.* 5(3):211–222.
18. Kaner, S., L. Lind, C. Toldi, S. Fisk, and D. Beger. 2011. *The facilitator's guide to participatory decision-making*. San Francisco, CA: Jossey-Bass. 368 p.
19. Management of conflict resolution & stimulation techniques. Available at: <https://theintactone.com/2019/03/14/ncm-u1-topic-7-management-of-conflict-resolution-stimulation-techniques/> (accessed February 26, 2021).
20. Tanuja, A. Conflict: Features, causes and consequences | Organisation. Available at: <https://www.businessmanagementideas.com/notes/management-notes/conflict-management/conflict-features-causes-and-consequences-organisation/5009> (accessed February 26, 2021).
21. Kolesnikov, A. V., and S. V. Listopad. 2018. Model' gibridnoy intellektual'noy mnogoagentnoy sistemy geterogennogo myshleniya dlya informatsionnoy podgotovki operativnykh resheniy v regional'nykh elektricheskikh setyakh [Model of a hybrid intelligent multiagent system of heterogeneous thinking for preparation of information about operational decisions in a regional power system]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(4):31–41. doi: 10.14357/08696527180404.

Received March 2, 2021

Contributors

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

Kirikov Igor A. (b. 1955) — Candidate of Sciences (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

СТРАТЕГИЯ ОБСЛЕДОВАНИЯ ПОМЕЩЕНИЙ ДЛЯ АВТОНОМНОГО МОБИЛЬНОГО РОБОТА НА ОСНОВЕ КАРТЫ РАКУРСОВ

О. А. Яковлев¹, П. О. Архипов²

Аннотация: Рассмотрен вариант задачи автономного обследования, цель которого — построение трехмерной реконструкции по снимкам, получаемым в процессе перемещения мобильного робота, оснащенного камерой. Предложена структура данных — *карта ракурсов*, накапливающая информацию о том, под какими углами наблюдались разные фрагменты сцены. Сформулирована стратегия обследования на основе карты ракурсов и описана ее эффективная реализация с применением частичных сумм и оптимального разбиения ортогонального многоугольника на прямоугольники. Также приведены результаты эксперимента в виртуальном окружении, подтверждающие возможность применения описанной стратегии в реальном времени. Оценена полнота полученной трехмерной модели помещения путем сравнения с эталоном.

Ключевые слова: трехмерная реконструкция; стратегия обследования; автономный мобильный робот

DOI: 10.14357/08696527210206

1 Введение

Обследование помещений — это инкрементальный процесс, на каждой итерации которого выполняется следующая последовательность процедур:

- (1) локализация робота;
- (2) обновление трехмерной модели путем интеграции очередного снимка;
- (3) выбор целевой позиции;
- (4) построение маршрута до целевой точки с использованием текущей модели помещения для обхода препятствий.

За выбор целевой позиции отвечает стратегия обследования — алгоритм, определяющий позицию робота, позволяющую максимально расширить представление об обследуемом окружении.

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

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

Среди существующих стратегий обследования можно выделить три большие группы: теоретико-информационные стратегии [1, 2], граничные стратегии [3–5] и стратегии наилучшей позиции съемки (next-best-view, NBV) [6, 7].

Теоретико-информационные стратегии рассматривают задачу обследования с позиции минимизации информационной энтропии. Стратегии этого класса обеспечивают качественный результат как с точки зрения точности реконструкции, так и с точки зрения общей длины маршрута обследования. Однако они имеют высокую вычислительную сложность, что исключает их применение в условиях реального времени.

Граничные стратегии базируются на разделении регионов области обследования на *проходимые*, *непроходимые* и *необследованные*. На каждой итерации процесса обследования целевой позицией робота назначается одна из границ между проходимым и необследованным регионами. Граничные стратегии нацелены главным образом на скорейшее построение навигационной карты робота, что не равносильно формированию точной трехмерной модели.

В NBV-стратегиях целью робота является позиция, обеспечивающая наилучшее дополнение или уточнение текущей трехмерной модели. Этот подход дает наиболее точную реконструкцию, однако не обеспечивает оптимальность маршрута.

В настоящей работе рассмотрена стратегия, принадлежащая классу NBV. Ключевая идея предложенной стратегии состоит в формировании структуры данных, основанной на принципе множественности наблюдений: необходимым условием реконструкции отдельно взятой точки ставится съемка с двух разных позиций, а для построения точной трехмерной модели необходима съемка одних и тех же фрагментов сцены под разными углами.

Рассмотрим распространенный вариант задачи обследования, в котором субъектом выступает робот с фиксированной камерой, перемещающийся по плоскости. В этом случае задача планирования перемещений робота становится двумерной. Примем, что оптическая система робота удовлетворяет модели проективной камеры [8], следовательно, имеет фиксированный горизонтальный угол обзора. Ограничим область перемещения робота прямоугольником, заведомо включающим обследуемое помещение, и разобьем его на ячейки равномерной решеткой, получив дискретную область обследования.

2 Карта ракурсов

По условию задачи камера неподвижна в системе координат робота, поэтому для описания его пространственного положения будем использовать тройку (x_c, y_c, φ) , где (x_c, y_c) — координаты проекционного центра камеры; φ — направление обзора, задаваемое углом относительно оси Ox . Пусть γ — горизонтальный угол обзора камеры. Тогда поле зрения камеры формируют такие точки (x, y) , что

$$2\text{angd}(\varphi, \text{atan2}(y - y_c, x - x_c)) \leq \gamma,$$

где

$$\text{angd}(\alpha, \beta) = \min(|\alpha - \beta|, 2\pi - |\alpha - \beta|);$$

$$\text{atan2}(y, x) = \begin{cases} 2 \arctg \frac{y}{\sqrt{x^2 + y^2} + x}, & x > 0 \vee y \neq 0; \\ \pi, & x < 0 \wedge y = 0. \end{cases}$$

Далее множество точек области обследования, попадающих в поле зрения камеры при положении робота (x_c, y_c, φ) , будем обозначать через $\text{fov}(x_c, y_c, \varphi)$.

Углом наблюдения точки (x, y) с позиции робота (x_c, y_c, φ) назовем величину $\theta(x, y, x_c, y_c)$ — значение полярного угла точки (x, y) относительно (x_c, y_c) . В примере, изображенном на рис. 1, этот угол выделен жирной линией.

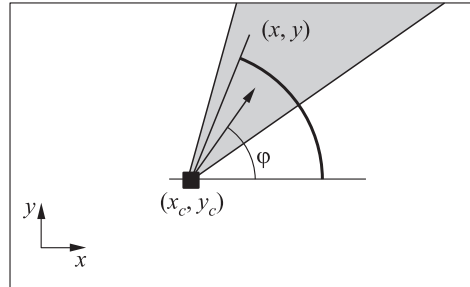


Рис. 1 Угол наблюдения точки

Разобьем область обследования на ячейки равномерной решеткой и рассмотрим последовательность снимков, сделанных с позиций (x_1, y_1, φ_1) , $(x_2, y_2, \varphi_2), \dots, (x_n, y_n, \varphi_n)$. Для каждой ячейки (x, y) эта последовательность порождает множество углов наблюдения:

$$\{\theta(x, y, x_i, y_i) | (x, y) \in \text{fov}(x_i, y_i, \varphi_i), i \in 1, 2, \dots, n\}.$$

Множества углов наблюдения для всех ячеек области обследования формируют карту углов наблюдения. Пример такой карты для двух позиций съемки приведен на рис. 2. Карта углов наблюдения дает представление о полноте реконструкции, а равно и полноте обследования, но ее анализ требует значительных вычислительных ресурсов, что затрудняет ее применение в условиях реального времени.

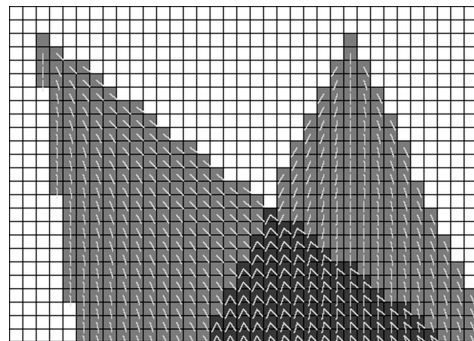


Рис. 2 Карта углов наблюдения

Введем приближение, заменив угол $\theta(x, y, x_c, y_c)$ углом ориентации робота φ . Таким образом, все точки, наблюдаемые роботом с некоторой позиции, будут иметь один и тот же угол наблюдения. Это свойство используется для

эффективной реализации процедуры выбора целевой точки, описанной в следующем разделе. Смысл такого приближения заключается в том, что угол ориентации робота — это средний угол наблюдения среди всех точек, попадающих в поле зрения, если не ограничивать поле зрения областью обследования. По определению поля зрения имеем:

$$|\varphi - \theta(x, y, x_c, y_c)| \leq \frac{\gamma}{2}.$$

Переход от карты углов наблюдения к карте ракурсов выполняется путем дискретизации угла с шагом $\pi/4$. Скажем, что точка наблюдалась в ракурсе $r \in \{1, 2, \dots, 8\}$, если она наблюдалась под таким углом α , что $\alpha \in [(r-1)\pi/4; r\pi/4)$. Таким образом, углу наблюдения α соответствует ракурс $\lfloor 4\alpha/\pi \rfloor + 1$. Карта ракурсов M каждой ячейке области обследования (x, y) ставит в соответствие вектор $M_{x,y} \in \{0, 1\}^8$, причем

$$M_{x,y,r} = \begin{cases} 1, & \text{если ячейка наблюдалась в ракурсе } r; \\ 0 & \text{в обратном случае.} \end{cases}$$

3 Стратегия обследования

По карте ракурсов M для всякого положения робота (x_c, y_c, φ) можно вычислить число ячеек области обследования, ранее не наблюдавшихся в ракурсе $r = \lfloor 4\varphi/\pi \rfloor + 1$:

$$\sum_{(x,y) \in \text{fov}(x_c, y_c, \varphi)} 1 - M_{x,y,r}. \quad (1)$$

Стратегию обследования можно сформулировать следующим образом: на очередной итерации процесса обследования среди позиций робота, достижимых из текущей, выберем ту позицию, которая максимизирует величину (1). Прямолинейная реализация этой стратегии, в которой для каждой позиции-кандидата (x_c, y_c, φ) просматривается все множество $\text{fov}(x_c, y_c, \varphi)$, неприменима в реальном времени, поэтому рассмотрим эффективную реализацию, основанную на принципе частичных сумм.

Матрицу $P = (p_{ij})_{n \times m}$ будем называть матрицей частичных сумм для матрицы $A = (a_{ij})_{n \times m}$, если $p_{kl} = \sum_{i=1}^k \sum_{j=1}^l a_{ij}$. При известной матрице P можно вычислять суммы на любой подматрице матрицы A , а именно: сумма подматрицы, содержащей элементы с индексами $i \in [a, b]$, $j \in [c, d]$, равна

$$p_{bd} - p_{bc-1} - p_{a-1d} + p_{a-1c-1}. \quad (2)$$

Чтобы выражение (2) было определено при $a = 1$ или $c = 1$, примем, что $p_{i0} = p_{0j} = 0$. Таким образом, вычислительная сложность расчета суммы

подматрицы с использованием частичных сумм составляет $O(1)$. В свою очередь, расчет матрицы P выполняется за линейное время согласно выражению:

$$p_{ij} = \begin{cases} a_{ij}, & i = 1, j = 1; \\ p_{ij-1} + a_{ij}, & i = 1; \\ p_{i-1j} + a_{ij}, & j = 1; \\ p_{i-1j} + p_{ij-1} - p_{i-1j-1} + a_{ij}, & i > 1, j > 1. \end{cases}$$

Заметим, что дискретное поле зрения камеры представляет собой ортогональный многоугольник, т. е. многоугольник со сторонами, параллельными координатным осям. Всякий ортогональный многоугольник можно разбить на прямоугольники со сторонами, параллельными осям координат (рис. 3).

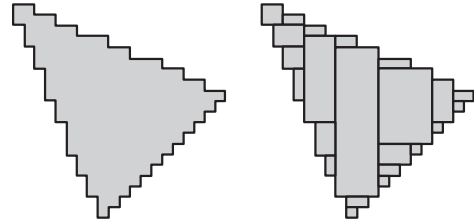


Рис. 3 Декомпозиция ортогонального многоугольника

Имея разбиение поля зрения на прямоугольники, можно вычислить величину (1), используя частичные суммы. Пусть $P_k = (p_{ij}^{(r)})$ — матрица частичных сумм для ракурса r ; n — число прямоугольников в разбиении; (a_i, b_i) — координаты верхней левой вершины i -го прямоугольника; (c_i, d_i) — координаты нижней правой вершины. Тогда с учетом (2) выражение (1) примет вид:

$$\sum_{i=1}^n p_{c_i d_i}^{(r)} - p_{c_i b_i-1}^{(r)} - p_{a_i-1 d_i}^{(r)} + p_{a_i-1 b_i-1}^{(r)}.$$

Таким образом, с использованием разбиения поля зрения на n прямоугольников величину (1) можно вычислить за время порядка $O(n)$, предварительно рассчитав матрицы частичных сумм. Следовательно, для ускорения процедуры выбора целевой позиции необходимо минимизировать число прямоугольников в разбиении.

Задача о разбиении ортогонального многоугольника на минимальное число прямоугольников сводится к задаче поиска максимального независимого множества в двудольном графе [9], которая по теореме Кёнига сводится к поиску наибольшего паросочетания. Поиск наибольшего паросочетания в двудольном графе можно выполнить с помощью алгоритма Хопкрофта–Карпа [10].

4 Эксперимент

Экспериментальная оценка предложенной стратегии обследования помещений выполнена в виртуальном окружении, реализуемом программным средством

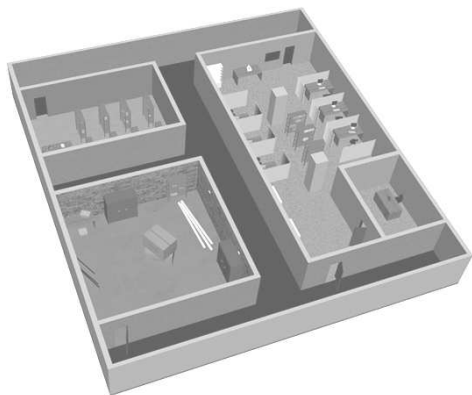


Рис. 4 Виртуальное окружение

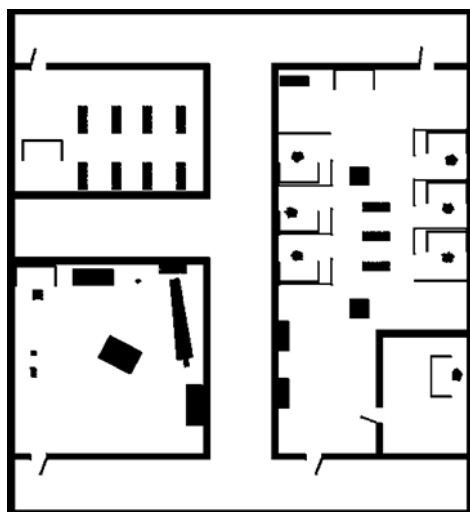


Рис. 5 Карта препятствий для поиска пути

шуму число ракурсов. Можно заметить, что полученный маршрут обследования не оптимален с точки зрения порядка обхода отдельных комнат, что объясняется локальной природой стратегии обследования. Среднее время выполнения процедуры поиска целевой позиции составило 6 мс, что подтверждает возможность применения рассматриваемого метода в реальном времени.

В результате обследования была построена трехмерная реконструкция помещения в виде множества точек, изображенного на рис. 7. Реконструкция получена

SfmDataGenerator [11]. Субъектом обследования в данном случае выступает виртуальный наблюдатель, перемещаемый в пределах сцены, а последовательность снимков для трехмерной реконструкции получается путем рендеринга сцены с позиции виртуального наблюдателя. В рамках эксперимента виртуальное окружение представлено трехмерной моделью помещения, изображенной на рис. 4. Для построения корректных маршрутов перемещения робота используется карта препятствий, полученная путем проекции на базовую плоскость области сцены, соответствующей высоте робота (рис. 5). Карта препятствий и область обследования имеют размер 513×558 ячеек. Поиск кратчайшего пути до целевой позиции, выбранной согласно предлагаемой стратегии обследования, выполняется с помощью алгоритма A^* [12].

Аппаратной частью эксперимента выступает персональный компьютер на базе процессора Intel Core i7 8700 с частотой 3,2 ГГц и 16 ГБ оперативной памяти. Графический процессор для вычислений не задействовался.

Ход эксперимента в виде маршрута обследования и карт ракурсов на отдельных итерациях показан на рис. 6. Карты ракурсов изображены в градациях серого таким образом, что светлые тона соответствуют боль-

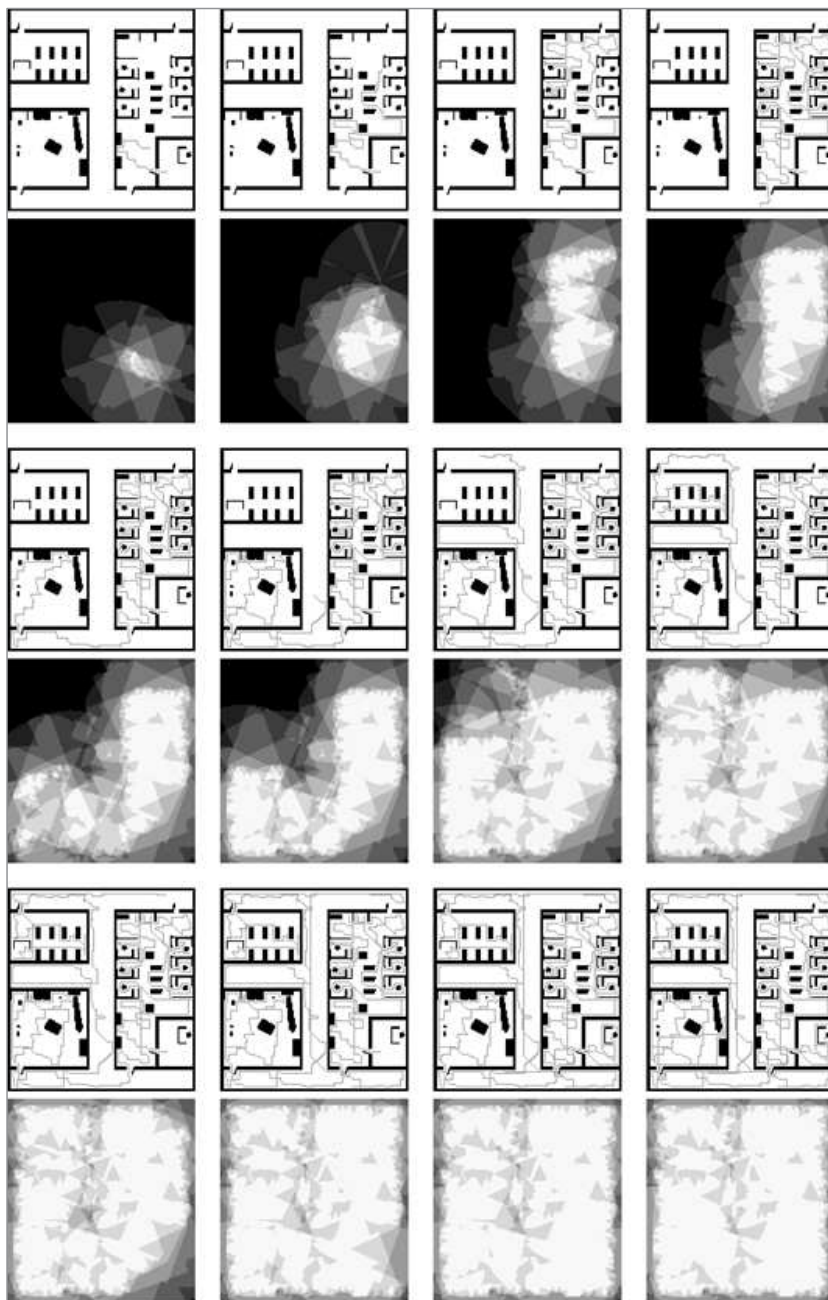


Рис. 6 Процесс обследования

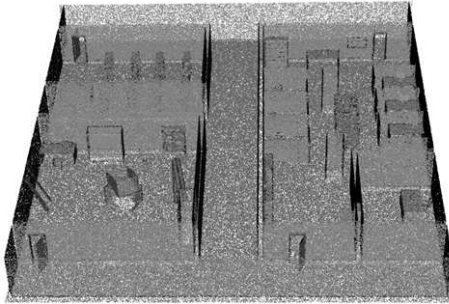


Рис. 7 Трехмерная реконструкция

и будем считать всякую точку $g \in G$ покрытой, если δ -окрестность этой точки содержит хотя бы одну точку $r \in R$. В свою очередь, показателем полноты обследования будем считать отношение числа покрытых точек к общему числу точек эталона $|G|$, выраженное в процентах: $\sum_{g \in G} [\exists r \in R : \|r - g\|_2 \leq \delta] / |G| \cdot 100\%$, где $[\cdot]$ — скобки Айверсона.

При $\delta = 2$ см полнота построенной реконструкции равна 97,54%.

5 Заключение

Представленная в статье структура данных позволяет осуществлять планирование перемещений автономного мобильного робота при обследовании помещений. Предложенная реализация стратегии обследования на основе карты ракурсов может применяться в приложениях реального времени, что подтверждено результатами измерения быстродействия. Результаты экспериментов показали, что рассмотренная стратегия не позволяет минимизировать общую протяженность маршрута, однако гарантирует достаточную полноту трехмерной реконструкции.

В качестве направления дальнейших исследований авторы рассматривают трехмерное обобщение карты ракурсов, позволяющее управлять роботами с большим числом степеней свободы, например, беспилотным летательным аппаратом.

Литература

1. Carrillo H., Dames P., Kumar V., Castellanos J. A. Autonomous robotic exploration using occupancy grid maps and graph SLAM based on Shannon and Rényi entropy // IEEE Conference (International) on Robotics and Automation Proceedings. — Piscataway, NJ, USA: IEEE, 2015. P. 487–494.
2. Li P., Yang C., Wang R., Wang S. A high-efficiency, information-based exploration path planning method for active simultaneous localization and mapping // Int. J. Adv. Robot. Syst., 2020. Vol. 17. Iss. 1. doi: 10.1177/1729881420903207.

путем обратной проекции карт глубины, сгенерированных SfmDataGenerator для каждого кадра.

Применение метода виртуальной съемки для апробации предложенной стратегии дает возможность формально оценить полноту обследования путем сравнения трехмерной реконструкции с эталоном. Для этого воспользуемся показателем на основе расстояния Хаусдорфа. Рассмотрим множество точек эталона G и множество точек реконструкции R . Зафиксируем допустимую точность δ

3. *Yamauchi B.* A frontier-based approach for autonomous exploration // IEEE Symposium (International) on Computational Intelligence in Robotics and Automation Proceedings. — Piscataway, NJ, USA: IEEE, 1997. P. 146–151.
4. *Архипов О. П., Гасилов А. В., Маньяков Ю. А., Яковлев О. А.* Алгоритм обследования замкнутого помещения автономным мобильным роботом // Системы и средства информатики, 2018. Т. 28. № 4. С. 10–21.
5. *Quin P., Nguyen D. D. K., Vu T. L., Alempijevic A., Paul G.* Approaches for efficiently detecting frontier cells in robotics exploration // Front. Robot. AI, 2021. Vol. 8. doi: 10.3389/frobt.2021.616470.
6. *Bircher A., Kamel M., Alexis K., Oleynikova H., Siegwart R.* Receding horizon “Next-Best-View” planner for 3D exploration // IEEE Conference (International) on Robotics and Automation Proceedings. — Piscataway, NJ, USA: IEEE, 2016. P. 1462–1468.
7. *Hardouin G., Moras J., Morbidi F., Marzat J., Mouaddib E. M.* Next-best-view planning for surface reconstruction of large-scale 3D environments with multiple UAVs // IEEE/RSJ Conference (International) on Intelligent Robots and Systems Proceedings. — Piscataway, NJ, USA: IEEE, 2020. P. 1567–1574.
8. *Hartley R., Zisserman A.* Multiple view geometry in computer vision. — 2nd ed. — Cambridge: Cambridge University Press, 2003. 670 p.
9. *Eppstein D.* Graph-theoretic solutions to computational geometry problems // Graph-theoretic concepts in computer science / Eds. C. Paul, M. Habib. — Lecture notes in computer science ser. — Springer, 2010. Vol. 5911. P. 1–16.
10. *Hopcroft J. E., Karp R. M.* An $n^{5/2}$ algorithm for maximum matchings in bipartite graphs // SIAM J. Comput., 1973. Vol. 2. Iss. 4. P. 225–231.
11. *Яковлев О. А., Гасилов А. В.* Создание реалистичных наборов данных для алгоритмов трехмерной реконструкции с помощью виртуальной съемки компьютерной модели // Системы и средства информатики, 2016. Т. 26. № 2. С. 98–107.
12. *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 15.03.21

INDOOR EXPLORATION STRATEGY FOR AUTONOMOUS MOBILE ROBOT BASED ON DIRECTION MAP

O. A. Yakovlev and P. O. Arkhipov

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

Abstract: The paper focuses on the variant of autonomous exploration problem that implies three-dimensional (3D) reconstruction of an unknown environment by moving a mobile robot with optical sensor through the environment. The authors introduce a data structure called direction map that accumulates information about viewpoints for every region of the scene. A direction map is used

to formulate an exploration strategy according to the next-best-view principle. This strategy can be efficiently implemented using rectangular partition of an orthogonal polygon and integral matrices. Experimental evaluation of the proposed strategy is done in virtual environment. Completeness of the resulting 3D model is evaluated by comparison with the ground truth.

Keywords: 3D reconstruction; indoor exploration; autonomous mobile robot

DOI: 10.14357/08696527210206

References

1. Carrillo, H., P. Dames, V. Kumar, and J. A. Castellanos. 2015. Autonomous robotic exploration using occupancy grid maps and graph SLAM based on Shannon and Rényi entropy. *IEEE Conference (International) on Robotics and Automation Proceedings*. Piscataway, NJ: IEEE. 487–494.
2. Li, P., C. Yang, R. Wang, and S. Wang. 2020. A high-efficiency, information-based exploration path planning method for active simultaneous localization and mapping. *Int. J. Adv. Robot. Syst.* 17(1). doi: 10.1177/1729881420903207.
3. Yamauchi, B. 1997. A frontier-based approach for autonomous exploration. *IEEE Symposium (International) on Computational Intelligence in Robotics and Automation Proceedings*. Piscataway, NJ: IEEE. 146–151.
4. Arkhipov, O. P., A. V. Gasilov, Yu. A. Maniakov, and O. A. Yakovlev. 2018. Algoritm obsledovaniya zamknutogo pomescheniya avtonomnym mobil'nym robotom [Enclosed room exploration algorithm for an autonomous mobile robot]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(4):10–21.
5. Quin P., D. D. K. Nguyen, T. L. Vu, A. Alempijevic, and G. Paul. 2021. Approaches for efficiently detecting frontier cells in robotics exploration. *Front. Robot. AI* 8. doi: 10.3389/frobt.2021.616470.
6. Bircher, A., M. Kamel, K. Alexis, H. Oleynikova, and R. Siegwart. 2016. Receding horizon “Next-Best-View” planner for 3D exploration. *Conference (International) on Robotics and Automation Proceedings*. Piscataway, NJ: IEEE. 1462–1468.
7. Hardouin, G., J. Moras, F. Morbidi, J. Marzat, and E. M. Mouaddib. 2020. Next-best-view planning for surface reconstruction of large-scale 3D environments with multiple UAVs. *IEEE/RSJ Conference (International) on Intelligent Robots and Systems Proceedings*. Piscataway, NJ: IEEE. 1567–1574.
8. Hartley, R., and A. Zisserman. 2003. *Multiple view geometry in computer vision*. 2nd ed. Cambridge: Cambridge University Press. 670 p.
9. Eppstein, D. 2010. Graph-theoretic solutions to computational geometry problems. *Graph-theoretic concepts in computer science*. Eds. C. Paul and M. Habib. Lecture notes in computer science ser. Springer. 5911:1–16.
10. Hopcroft, J. E., and R. M. Karp. 1973. An $n^{5/2}$ algorithm for maximum matchings in bipartite graphs. *SIAM J. Comput.* 2(4):225–231.
11. Yakovlev, O. A., and A. V. Gasilov. 2016. Sozdanie realistichnykh naborov danykh dlya algoritmov trekhmernoy rekonstruktsii s pomoshch'yu virtual'noy s'emki komp'yuternoy modeli [Generating realistic structure-from-motion datasets through

virtual photography]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 26(2):98–107.

12. Hart, P. E., N. J. Nilsson, and B. Raphael. 1968. Formal basis for the heuristic determination of minimum cost paths. *IEEE T. Syst. Sci. Cyb.* 4(2):100–107.

Received March 15, 2021

Contributors

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

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

НЕКОТОРЫЕ ВОПРОСЫ ПРОГРАММНОГО ОПРЕДЕЛЕНИЯ ХРАНИЛИЩА ДАННЫХ

*В. Б. Егоров*¹

Аннотация: Развитие информационных технологий (information technologies, IT) вовлекает в сферу интересов владельцев центров обработки данных (ЦОД) наряду с более совершенным оборудованием хранилищ данных также и более эффективные технологии их хранения. Одной из таких технологий обещает стать программно определяемое хранилище (software-defined storage, SDS). Следуя за программным определением сетей (software-defined networking, SDN), SDS заимствовало у лидера ряд базовых положений, в первую очередь реализацию плоскости управления хранением данных в едином программном контроллере и упрощение оборудования плоскости данных. Однако при растущем интересе к SDS в его теоретическом базисе все еще сохраняются определенные пробелы, начиная с отсутствия для него общепринятого адекватного определения. Нередко эти пробелы заполняются в посвященной SDS литературе сомнительными и некорректными утверждениями. Отсутствует и понимание объективных различий между SDS и SDN, заставляющих эти две родственные технологии развиваться несколько разными путями. В статье сделана попытка дать «универсальное» определение SDS, выявлены принципиальные отличия SDS от SDN, показаны особенности развития этой перспективной технологии.

Ключевые слова: гиперконвергентная инфраструктура; «оркестровка»; программно определяемое хранилище

DOI: 10.14357/08696527210207

1 Введение

Бурное развитие информационных технологий вовлекает в обработку и хранение все большие объемы данных. Соответственно растет значение осуществляющих эти функции технических средств и программного обеспечения (ПО) в ЦОД. Необходимость удовлетворять растущие запросы все большего числа пользователей привела к появлению в ЦОД, в частности самых разнообразных систем хранения данных (СХД), начиная от простых непосредственно подключаемых (directly-attached storage, DAS) до наделенных интеллектом удаленных: сетевых хранилищ (network-attached storage, NAS) и сетей хранения данных (storage-area network, SAN). Однако качественное усложнение задач, встающих перед IT-системами с появлением «больших данных» (big data) и «облачных»

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

(cloud) технологий, начинает вовлекать в сферу интересов владельцев ЦОД помимо более совершенного оборудования также и более эффективные подходы к хранению данных. Одним из таких перспективных подходов обещает стать программное определение хранилищ данных.

Концепция программно определяемого хранилища возникла не на пустом месте. Родившись на волне интереса к программно определяемым сетям, она переняла у предшественницы ряд ключевых понятий и положений. Но, не довольствуясь ролью ведомого, SDS уже уверенно обходит SDN по числу примеров практических воплощений [1]. Тем не менее при всех своих успехах (и, вероятно, во многом благодаря им) SDS сегодня выглядит скорее пестрой совокупностью более или менее удачных маркетинговых решений, нередко объединяемых лишь претензией на программное определение. Потенциальному пользователю в этом нагромождении непросто отделить зерна от плевел, в том числе и вследствие отсутствия надежных теоретических основ программного определения хранилищ. Ниже сделана попытка восполнить некоторые пробелы в этих основах и показать принципиальные особенности, отличающие SDS от SDN.

2 Определение и сущность программно определяемого хранилища

Теоретические проблемы SDS начинаются с отсутствия в IT-индустрии его адекватного общепринятого определения. Разные поставщики продуктов с претензией на причастность к SDS, причем как аппаратуры, так и ПО, предлагают свои дефиниции [2]. Определение корпорации IDC, например, включает в SDS «любой стек ПО хранения, который может быть установлен на типовом оборудовании и способен предложить полный набор сервисов и интеграцию наличных ресурсов хранения для обеспечения мобильности данных между этими ресурсами». Корпорация Dell EMC определяет SDS как «ПО хранения, не определяемое оборудованием, на котором оно работает, и способное быть загруженным и работать на любом типовом промышленном оборудовании». Для компании NetAPP SDS — это «программируемость ресурсов хранения, позволяющая абстрагировать их от физической аппаратной платформы для лучшей масштабируемости, большей гибкости и эффективности». Вряд ли из такого рода не очень согласуемых между собой дефиниций, во многом рекламно-прикладного плана, можно составить представление о сущности предмета. Не улучшает ситуацию определение законодателя в области СХД — Ассоциации промышленности сетевых хранилищ (Storage Networking Industry Association, SNIA): «Виртуализированное хранилище с интерфейсом управления услугами, включающим автоматизацию, виртуализированный тракт данных, масштабируемость и прозрачность». Наконец, как и в случае SDN [3], наиболее информативной оказывается Wikipedia, определяющая SDS как автоматизированную инфраструктуру хранения данных, которая управляется интеллигентным ПО, эффективно подстраивается под требования приложений и расширяет функциональные возможности СХД добавлением таких функций, как компрессия данных, «тонкая настройка» (thin

provisioning), дедупликация, репликация, «моментальные снимки» (snapshots) и др. Такое определение SDS не только наиболее адекватно и информативно, но и вполне подпадает под универсальное «практичное» определение программного определения как комплексной автоматизации централизованного управления инфраструктурой ЦОД специализированным ПО на основе заданных политик [4].

Следует заметить, что у некоторых поставщиков SDS-продуктов программное определение хранилища оказывается практически синонимом его виртуализации, что видно, в частности, из приведенных выше дефиниций NetAPP или SNIA. Но SDS не то же самое, что виртуализация хранилища. Виртуализация представляет имеющиеся разрозненные устройства хранения в виде единого пула ресурсов, т. е. абстрагирует наличные и доступные для использования объемы хранения от физического оборудования, предоставляющего эти объемы. А SDS идет дальше, абстрагируя от наличного физического оборудования не только объемы, но также сервисы хранения и характеристики хранилища [5], т. е. виртуализацию можно рассматривать как компонент, но не эквивалент SDS.

3 Цели, средства и решаемые задачи

Главная цель внедрения SDS — более эффективное использование ресурсов хранения ЦОД и более полное удовлетворение возрастающих требований пользователей к хранению данных. Средства достижения этой цели — централизация управления инфраструктурой хранения и автоматизация менеджмента предоставления сервисов. Помимо улучшения использования ресурсов хранения централизация и автоматизация упрощают администратору ЦОД мониторинг и управление функционированием инфраструктуры хранения.

Ценное качество SDS — возможность масштабирования ресурсов хранения «расширением» (scaling-out) вместо их «наращивания» (scaling-up). Такое масштабирование позволяет увеличивать ресурсы хранения без добавления оборудования, заставляя имеющиеся СХД работать более эффективно, в частности за счет «тонкой настройки» и динамического выделения здесь и сейчас простаивающих ресурсов для временного хранения данных разных пользователей. Кроме того, абстрагирование ресурсов и сервисов хранения от аппаратуры СХД упрощает привлечение при необходимости «облачных» хранилищ.

Одним из наиболее рекламируемых свойств SDS является независимость оборудования СХД от поставщика, а аппаратуры и ПО — друг от друга [6]. Считается, что SDS позволяет продлевать жизнь устаревающих СХД, не только увеличивая их эффективные емкости за счет масштабирования «вширь», но и повышая надежность хранения в них благодаря возможности парировать последствия отказов отдельных единиц оборудования распределением объемов работ по всей инфраструктуре хранения данных [7].

Также предполагается, что вследствие виртуализации трактов данных SDS предоставляет пользователю возможность работы в равной мере с любыми типами интерфейсов: блочными, файловыми или объектными.

Итак, комплекс наиболее часто отмечаемых достоинств SDS включает:

- более эффективное использование имеющихся ресурсов хранения данных;
- дополнительные возможности масштабирования, теоретически неограниченные с привлечением «облачных» технологий;
- поддержка разнообразных интерфейсов доступа (блочного, файлового, объектного);
- возможность комплектации инфраструктуры ЦОД СХД оборудованием от разных производителей;
- централизованное управление всеми ресурсами хранения через единый административный интерфейс.

4 Сопоставление программно определяемых хранилищ и сетей

Поскольку концепция SDS родилась вслед за концепцией SDN и под ее прямым влиянием, неудивительно, что SDS и SDN преследуют во многом схожие цели и характеризуются рядом общих основополагающих черт. Однако объективная разница назначения и функций у сетей и хранилищ неизбежно приводит к некоторым различиям между SDN и SDS, не всегда бросающимся в глаза.

Общность SDS и SDN прежде всего проявляется в физическом разделении плоскостей данных (data plane) и управления (control plane) [8]. Физическое разделение плоскостей принципиально необходимо в качестве предпосылки централизации менеджмента, которую можно интерпретировать как стягивание плоскости управления в точку — логически единый программный контроллер, работающий на типовом сервере или виртуальной машине (VM). В SDN одним из положительных следствий централизации менеджмента предполагается упрощение плоскости данных, в частности полное исключение маршрутизаторов и сведение номенклатуры оборудования к простым унифицированным коммутаторам, управляемым из контроллера сети по единому протоколу типа OpenFlow. Аналогичные решения предлагаются и для SDS [9]. Но в случае SDS задача усложняется бóльшим разнообразием оборудования хранения данных и отсутствием протокола, аналогичного OpenFlow. Решение последней проблемы ожидалось от SNIA, но, кажется, оно может появиться не столько в виде унифицированного протокола, сколько в виде стандартизованных интерфейсов прикладного программирования.

В SDN концептуально предполагается, что ПО контроллера сети, нередко называемое «оркестровщиком», работает на прикладном уровне. При этом парадоксальным образом этому управляющему сетью ПО не требуется сетевая операционная система (ОС), так как контроллер SDN, управляя сетью, не является абонентом сети. Следовательно, «оркестровщик» может работать с любой стандартной ОС, а также в VM или даже контейнере (container). Но в отношении

SDS ситуация иная. В простейшем случае, когда ЦОД имеет единственное СХД, например типа NAS, в функции ПО SDS входит только отработка сервисов хранения, что при определенных допущениях реализуемо на прикладном уровне. Если же в ЦОД есть несколько разных СХД либо он включает современную гиперконвергентную инфраструктуру (hyper-converged infrastructure, HCI) с множеством DAS, распределенных по серверам и находящихся под контролем локальных ОС и гипервизоров, ПО SDS тоже вынуждено внедряться во все серверы ЦОД, связанные с инфраструктурой хранения, и затрагивать системное ПО вплоть до гипервизорного уровня. Следовательно, разработчики HCI и поставщики системного ПО для такой инфраструктуры должны так или иначе встраивать в свои продукты поддержку SDS. Правда, в таком случае в концептуальном плане нарушается предполагаемая «по умолчанию» реализуемость программного определения на прикладном уровне, а в практическом — оказывается под угрозой широко рекламируемая независимость ПО SDS от поставщика оборудования СХД.

Также различно отношение SDS и SDN к виртуализации. Если для SDN виртуализация не принципиальна [10], то реализация программного определения хранилища без консолидации всех ресурсов хранения в единый виртуальный пул выглядит крайне проблематичной. Не случайно, как было отмечено выше, у ряда поставщиков SDS-продуктов программное определение хранилища фактически отождествляется с его виртуализацией. Таким образом, хотя можно встретить и противоположное мнение [5], все же, по-видимому, для SDS, в отличие от SDN, виртуализацию следует полагать обязательным условием.

Наиболее же существенны различия между SDN и SDS в способах их развертывания.

5 Особенности развертывания программно определяемых хранилищ

В SDN наибольшую трудность представляет разработка и отладка ПО «оркестровки» для конкретной сети в конкретном ЦОД. После того как эта работа выполнена, развертывание SDN сводится к рутинной процедуре запуска «оркестровщика» в контроллере сети.

Программное обеспечение SDN монолитно и изолированно. «Оркестровщик» работает на отдельном сервере, никак не пересекаясь с другими серверами ЦОД и работающими на них программами — как системными, так и прикладными.

Программное обеспечение SDS в общем случае не монолитно и не изолированно. Оно вынужденно распадается на две функциональные части: SDS-контроллер, работающий аналогично SDN-контроллеру в качестве приложения на некоем сервере (либо в ВМ или контейнере), и некое множество SDS-агентов, внедренных в системное ПО прочих серверов ЦОД. Контроллер осуществляет общее централизованное управление и мониторинг работы СХД, а также обрабатывает запросы сервисов хранения от всех приложений в ЦОД. Задача SDS-агентов — перехватывать «на местах» эти запросы и перенаправлять их

в SDS-контроллер. Таким образом, ПО SDS не может ограничиться одним отдельным приложением или даже одним сервером, а должно стать составной частью системного ПО на всех серверах ЦОД, включающих какие-то ресурсы хранения данных.

Однако если обратиться к рекламе имеющихся на рынке SDS-продуктов, эту особенность развертывания SDS обнаружить непросто. Так, типичное для этого рынка предложение компании DataCore включает три варианта поставки и развертывания SDS [11]:

- (1) чисто программный (software-only) — ПО SDS, устанавливаемое у пользователя на типовом сервере x86;
- (2) с сервером (appliance) — сервер с предустановленным ПО SDS;
- (3) с сервером и хранилищем (appliance + storage) — интегрированная система с сервером, встроенным хранилищем и предустановленным ПО SDS.

Реклама обещает возможность работы поставляемого ПО SDS с любыми внутренними и внешними хранилищами от произвольных поставщиков, но только для двух первых вариантов поставки. В третьем варианте с конкретным встроенным хранилищем мультивендорная поддержка не предусматривается. Это означает, что ПО SDS все-таки требует настройки на конкретный тип СХД, т. е. ему требуется своеобразная «оркестровка». И, будучи «оркестрованным» под конкретную поставляемую в комплексе СХД, оно уже не в состоянии поддерживать другие типы хранилищ. Другое дело, что в отличие от «оркестровки» SDN, которая выполняется по месту под конкретную сеть и потому доступна весьма ограниченному кругу пользователей, SDS «оркеструется» под конкретный тип СХД и ПО SDS может быть приобретено пользователем в готовом виде, в частности от поставщика СХД вместе с оборудованием. Особенно ярко эта особенность SDS проявляется в HCI, где уже сегодня многие продукты включают некие предустановленные средства программного определения [12].

По всей видимости, именно эта особенность SDS, позволяющая пользователю приобрести на рынке готовый «оркестровщик» для конкретного типа СХД, объясняет тот факт, что SDS, будучи моложе SDN, уже вырвалось вперед по востребованности и числу работающих примеров.

6 Заключение

Концепция программного определения хранилища данных приобретает все большую популярность. Перенос рожденной для сетевой инфраструктуры ЦОД идеи программного определения на инфраструктуру хранения данных позволил заимствовать от SDN ряд понятий и базовых положений, в том числе центрацию плоскости управления в едином контроллере и снижение требований к аппаратуре плоскости данных. Тем не менее объективные различия в оборудовании и функциональности между сетями и хранилищами неизбежно привнесли

в SDS существенные особенности, не свойственные SDN. Эти особенности, похоже, не до конца осмыслены даже поставщиками оборудования и ПО SDS, что нашло отражение в отсутствии общепринятого определения предмета и ряде некорректных рекламно-маркетинговых утверждений.

В целом SDS вполне подпадает под универсальное определение программного определения как централизации управления инфраструктурой ЦОД и ее комплексной автоматизации с помощью специализированного ПО «оркестровки» на основе заданных политик. Главное же отличие SDS от SDN заключается в разветвлении и характере «оркестровки». В SDN она имеет объектом конкретную сеть, практически уникальную в каждом ЦОД, и потому, как правило, предполагает разработку и отладку по месту уникального «оркестровщика». В SDS «оркестровка» более локальна и имеет дело с массово выпускаемым оборудованием СХД, поэтому у владельца ЦОД появляется возможность приобретать готовый «оркестровщик» вместе с СХД от его поставщика или стороннего разработчика ПО. Благодаря этой приятной особенности SDS уже обошло SDN по числу работающих систем и, следует ожидать, продолжит лидировать с возрастающим отрывом.

С практической точки зрения интерес к SDS был спровоцирован с двух противоположных концов: с одной стороны возможностью продлить активную жизнь устаревающим СХД, а с другой — внедрением в ЦОД перспективных НСІ. Теоретически благодаря SDS замена устаревающего оборудования хранения данных может происходить постепенно и относительно безболезненно, не нарушая функционирования ЦОД, при этом смягчаются ограничения на типы нового оборудования и выбор его поставщиков. В последнее же время императивом обращения к SDS владельцев ЦОД становится необходимость полноценно использовать хранилищные ресурсы современных НСІ, в которых распыленным по множеству стереотипных серверов ресурсам хранения принципиально требуется централизованный консолидирующий и распределяющий эти ресурсы менеджер, в качестве которого органично смотрится SDS.

Понимание особенностей SDS, включая присущие ему отличия от SDN, должно способствовать еще более широкому распространению SDS в качестве как самостоятельной перспективной технологии, так и естественного компонента целиком программно определяемого ЦОД.

Литература

1. *Millman R.* Software-defined storage vs NAS/SAN: What are the options? // Computer Weekly, 2019. <https://www.computerweekly.com/feature/Software-defined-storage-vs-NAS-SAN-What-are-the-options>.
2. *Harvey C.* What is software defined storage? // Enterprise Storage, 2018. <https://www.enterprisestorageforum.com/storage-software/what-is-software-defined-storage.html>.

3. *Егоров В. Б.* Некоторые вопросы практической реализации концепции SDN // Системы и средства информатики, 2016. Т. 26. № 1. С. 109–120.
4. *Егоров В. Б.* Практичное определение «программно определяемого» // Системы и средства информатики, 2019. Т. 29. № 2. С. 85–94.
5. StarWind. What is software-defined storage? <https://www.starwindsoftware.com/software-defined-storage>.
6. Software-defined storage (SDS). TechTarget, 2017. <https://searchstorage.techtarget.com/definition/software-defined-storage>.
7. *Шуклин Г.* Рассуждения о Software Defined Storage: что не так с IO? // Хабр, 2014. <https://habr.com/ru/post/124755>.
8. *Jararweh Y., Al-Ayyoub M., Darabseh A., Benkhelifa E., Vouk M., Rindos A.* Software defined cloud: Survey, system and evaluation // Future Gener. Comp. Sy., 2016. Vol. 58. P. 56–74.
9. *O'Reilly J.* Software-defined storage // Network storage: Tools and technologies for storing your company's data. — Elsevier Inc., 2017. doi: B978-0-12-803863-5.00005-4.
10. *Егоров В. Б.* Некоторые неявные вопросы сетевой виртуализации // Системы и средства информатики, 2017. Т. 27. № 3. С. 88–98.
11. DataCore. What is software-defined storage? <https://www.datacore.com/software-defined-storage>.
12. *Edwards J.* Software-defined storage: Getting started // NetworkComputing, 2018. <https://www.networkcomputing.com/data-centers/software-defined-storage-getting-started>.

Поступила в редакцию 22.07.20

SOME ISSUES OF THE SOFTWARE-DEFINED STORAGE

V. B. Egorov

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

Abstract: Evolution of information technologies attracts the interest of data center owners, alongside the more functional storage equipment, to more efficient storing technologies, the software-defined storage (SDS) being among promising ones. Following the software-defined networking (SDN), the SDS has borrowed some essentials from the precursor, implementation of the control plane in a single programmable controller and simplification of the data plane hardware included. However, despite the growing interest to the SDS, some definite gaps remain in the SDS theoretical basics, beginning with the lack of a generally accepted

adequate definition. As a result, a lot of questionable and incorrect assertions may be found in publications devoted to the SDS. Also, the publications do not demonstrate comprehension of the objective distinction between the SDS and SDN, which compel these related technologies to evolve in somewhat differing ways. In the article, an attempt is made to formulate a “universal” definition of the SDS, principal distinctions of the SDS from the SDN are exposed, and some development features of this promising technology are outlined.

Keywords: hyper-converged infrastructure (HCI); orchestration; software-defined storage (SDS)

DOI: 10.14357/08696527210207

References

1. Software-defined storage vs NAS/SAN: What are the options? Available at: <https://www.computerweekly.com/feature/Software-defined-storage-vs-NAS-SAN-What-are-the-options> (accessed March 17, 2021).
2. What is software defined storage? Available at: <https://www.enterprisestorageforum.com/storage-software/what-is-software-defined-storage.html> (accessed March 17, 2021).
3. Egorov, V. B. 2016. Nekotorye voprosy prakticheskoy realizatsii kontseptsii SDN [Some issues of the SDN concept practical implementation]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 26(1):109–120.
4. Egorov, V. B. 2019. Praktichnoe opredelenie “programmno opredelyaemogo” [A practical definition of “software-defined”]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(2):85–94.
5. StarWind. What is software-defined storage? Available at: <https://www.starwindsoftware.com/software-defined-storage> (accessed March 17, 2021).
6. TechTarget. 2017. Software-defined storage (SDS). Available at: <https://searchstorage.techtarget.com/definition/software-defined-storage> (accessed March 17, 2021).
7. Shuklin, G. 2014. Rassuzhdeniya o Software Defined Storage: chto ne tak s IO? [Discourses on the software-defined storage: What is wrong with IO?]. Habr. Available at: <https://habr.com/ru/post/124755/> (accessed March 17, 2021).
8. Jararweh, Y., M. Al-Ayyoub, A. Darabseh, E. Benkhelifa, M. Vouk, and A. Rindos. 2016. Software defined cloud: Survey, system and evaluation. *Future Gener. Comp. Sy.* 58:56–74.
9. O’Reilly, J. 2017. Software-defined storage. *Network storage: Tools and technologies for storing your company’s data*. Elsevier Inc. 83–94. doi: B978-0-12-803863-5.00005-4.
10. Egorov, V. B. 2017. Nekotorye neyavnye voprosy setevoy virtualizatsii [Some implicit issues of network virtualization]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 27(3):88–98.

11. DataCore. What is software-defined storage? Available at: <https://www.datacore.com/software-defined-storage/> (accessed March 17, 2021).
12. Edwards, J. 2018. Software-defined storage: Getting started. NetworkComputing. Available at: <https://www.networkcomputing.com/data-centers/software-defined-storage-getting-started> (accessed March 17, 2021).

Received July 22, 2020

Contributor

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

ИНФОРМАЦИОННЫЕ АСПЕКТЫ ОБЕСПЕЧЕНИЯ БЕЗОПАСНОСТИ НА ТРАНСПОРТЕ: ПОИСК И ОТБОР ИНФОРМАЦИИ

А. В. Борисов¹, А. В. Босов², Д. В. Жуков³, А. В. Иванов⁴

Аннотация: Вторая статья цикла, посвященного описанию системы поддержки принятия решений (СППР) в области обеспечения безопасности пассажирских перевозок и противодействия противоправной деятельности на транспорте. На основе концептуальной модели, предложенной авторами в первой статье, формируется функциональное наполнение системы. Предложена классификация аналитических функций, включающая поисковые запросы, аналитические расчеты и задачи процессинга данных, и представлено содержание задач первого (поискового) типа, специализированных в рамках предметной области. Часть поисковых запросов, имеющих универсальное назначение, объединена в группу служебных методов. Вторую часть составили специализированные транспортные поисковые запросы. Содержательные в рамках предметной области поиски анонсируют запланированное представление более глубоких аналитических функций.

Ключевые слова: транспортная безопасность; система поддержки принятия решений; поисковый запрос; расчетная задача; процессинг данных

DOI: 10.14357/08696527210208

1 Введение

В работе [1], первой в цикле, который продолжает настоящая статья, инициировано исследование предметной области пассажирских перевозок с точки зрения организации информационного обеспечения в СППР, создаваемых в интересах обеспечения безопасности транспортных перевозок. Представленная в [1] концептуальная модель предметной области описывает как базовые понятия (пассажир, перевозчик, поездка, рейс, транспортный пункт), отражающие факт перевозки конкретного физического лица из одного транспортного пункта в другой, так и специфические (сложный пассажир, транспортная зона), не-

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

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

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

⁴Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, A.Ivanov@ipiran.ru

обходимые при выполнении аналитической обработки информации. Перечень атрибутов, конкретизирующих понятия модели, содержит ряд атрибутов технологического характера, роль которых становится ясна из детализации решаемых аналитических задач.

Создание концептуальной модели стало результатом проработки функционального наполнения СППР, основное назначение которой состоит в информационной и аналитической поддержке деятельности экспертов в сфере обеспечения безопасности на транспорте. Однако нет препятствий для более широкого применения такой СППР, например в решении задач противодействия противоправной деятельности.

Деятельность экспертов, использующих такую СППР, включает выполнение различных задач. Всего можно выделить три их типа. Первый и самый простой — поиски. Как и в любой аналитической работе, отбор исходных данных является важнейшим и необходимым этапом. В статье будет рассмотрено подмножество специфических запросов, характерных именно для пассажирских перевозок в связи с мероприятиями по обеспечению безопасности на транспорте или противодействию противоправной деятельности. Варианты задач, связанных с выполнением поисков, обсуждаются в третьем разделе статьи, применяемые вспомогательные/технологические процедуры — во втором.

В качестве второго этапа аналитики, второго типа задач предлагается некоторый круг постановок, уже не ограниченных просто отбором информации, но требующих выполнения некоторых расчетов с отобранными данными. Наконец, третий тип задач составляют сложные вычислительные процедуры, требующие значительных вычислительных ресурсов, в том числе на предварительную подготовку данных. Характерным средством решения такого рода задач служат инструменты интерактивного процессинга данных (OLAP, online analytical processing). Статья продолжает исследование в выбранной предметной области [1], представляя аналитические задачи первого типа. Расчетные задачи второго типа планируется представить в третьей статье цикла.

2 Служебные методы

Необходимым свойством СППР, реализующих аналитический функционал, должна быть способность работы с неточно заданной информацией. Любые текстовые атрибуты могут содержать ошибки, которые возникают как на этапе первоначального сбора информации (в рассматриваемой области — транспортной), так и в ходе получения оперативной информации, используемой затем при выполнении поисковых запросов. Необходимость работы с неточно заданной информацией требует как соответствующей алгоритмической поддержки, так и наличия определенной стратегии в работе с такой информацией. Непосредственное применение нечетких алгоритмов поиска на неподготовленных данных может быть медленным. Поскольку максимальное время выполнения поисковых запросов ограничено, уже на этапе проектирования информационной

модели целесообразно наряду с сохранением исходных атрибутов предусмотреть их преобразование и сохранение в виде, ускоряющем выполнение поисковых алгоритмов.

Примером такого преобразования служит построение фонетических ключей. Чтобы сократить время поиска, фонетические ключи для фамильно-именной группы атрибутов формируются заблаговременно, в процессе наполнения хранилища данных.

Ниже перечислены служебные функции, обеспечивающие поддержку работы с неточно заданной информацией:

- (1) построение русских и английских фонетических ключей;
- (2) поиск нестрогого совпадения строк с заданным числом ошибок;
- (3) определение близости номеров мест;
- (4) перевод из кириллицы в латиницу и обратно;
- (5) сравнение двух пассажиров (идентификация и поиск похожих пассажиров);
- (6) проверка совместности поездок пассажиров.

Для построения русских фонетических ключей апробировался и показал удовлетворительные результаты алгоритм «Русский Metaphone» [2], преобразующий исходные слова на русском языке с учетом фонетического звучания безударных гласных и возможных «слияний» согласных при произношении. Для построения английских фонетических ключей — алгоритм «Metaphone» [3], преобразующий исходное слово с учетом правил фонетики английского языка.

Для нестрогого сравнения двух строк S_1 и S_2 (фамилия, имя или отчество) известно множество алгоритмов [4]. В работе применялось сравнение, основанное на вычислении «расстояния Левенштейна» $d(S_1, S_2)$ [5]. Если $d(S_1, S_2) > r$, где r — настраиваемый параметр, равный допустимому числу ошибок и задаваемый в запросе пользователем, то строки S_1 и S_2 считаются сильно различающимися, иначе — близкими.

Текстовые атрибуты могут быть представлены как в кириллице, так и в латинице. Для перевода из латиницы в кириллицу фамильно-именной группы используются табличные функции перекодировки.

В запросах, требующих определения степени похожести пассажиров, использовалась скалярная функция сравнения данных пассажиров. Данная функция получает в качестве параметров номер документа, ФИО, пол, дату рождения для сравниваемых пассажиров и вычисляет общую степень сходства как сумму степеней сходства входных параметров, взятых с заранее заданными весовыми коэффициентами. Степень сходства текстовых параметров (номер документа, ФИО) определяется путем вычисления расстояния Левенштейна. Степень сходства полов определяется функцией, возвращающей 0, если полы не совпадают, 0,5, если один из аргументов не определен, и 1 — в остальных случаях. Степень сходства дат определяется числом несовпадающих символов в записи даты.

В запросах, требующих определения совместного проезда нескольких пассажиров, используется скалярная функция проверки совместности поездок. Эта функция получает в качестве параметров два идентификатора поездок, тип алгоритма поиска, флаг проверки проезда на близко расположенных местах и вычисляет булево значение, определяющее факт совместного нахождения пассажиров на маршруте поездки. Возможны три алгоритма работы функции:

- (1) поездки считаются совместными, если совпадают начальный и конечный пункты и даты нахождения в них;
- (2) поездки считаются совместными, если конечный пункт одной поездки совпадает с начальным пунктом другой поездки и даты нахождения в них совпадают;
- (3) поездки считаются совместными, если пассажиры следовали одним маршрутом.

3 Транспортные поисковые запросы

Применительно к рассматриваемой предметной области оперативная работа или действия по расследованию инцидентов могут быть отнесены к разновидности аналитической деятельности. Первый этап такой деятельности — отбор исходных данных [6]. На этом этапе необходимо предоставить пользователю системы набор поисковых запросов, позволяющий получить как базовую информацию о пассажирах и их поездках, так и установить наличие некоторых взаимосвязей между пассажирами. Кроме того, могут быть полезны и запросы, позволяющие получить данные о транспортных средствах, о регистрации пассажиров, а также проверить данные на непротиворечивость.

Для поисковых запросов введена простая классификация:

- (1) типовые фильтры;
- (2) задачи поиска и объединения пассажиров;
- (3) специализированные запросы по авто- и авиаперевозкам;
- (4) поиск артефактов, несоответствий, различий в данных о поездках.

К группе *типовых фильтров* относятся:

- (1) поиск пассажиров и их поездок по точно заданным значениям параметров, обеспечивающий отбор пассажиров по атрибутам, заданным в классе «Пассажир»; для каждого пассажира, включенного в результат выполнения запроса, определяется список его поездок;
- (2) поиск пассажиров и их поездок с учетом фонетических искажений; данный запрос аналогичен предыдущему, но позволяет включить в результат выполнения запроса пассажиров с учетом возможных фонетических искажений их текстовых атрибутов: например, в запросе задана фамилия «Агапонов», а в результат будет включен пассажир с фамилией «Агафонов»;

- (3) поиск пассажиров с возможными ошибками; данный запрос аналогичен первому запросу, но позволяет выполнить поиск с нестрогим совпадением строк;
- (4) поиск пассажиров при заданных атрибутах в кириллице или латинице; данный запрос аналогичен первому запросу, но позволяет выполнять поиск, если исходные данные пассажира хранятся в латинице, а в поисковой форме заданы в кириллице и наоборот;
- (5) поиск «похожих» пассажиров; данный запрос позволяет для группы заранее заданных атрибутов пассажиров выполнить поиск «похожих» пассажиров, т. е. пассажиров, для которых степень сходства атрибутов с заранее заданными превысит заданное оператором значение; степень сходства вычисляется с помощью скалярной функции сравнения данных пассажиров, описанной в разд. 2;
- (6) поиск сведений о поездках, позволяющий для заданного пассажира или группы пассажиров выполнить поиск данных о совершенных ими поездках;
- (7) поиск рейсов для транспортного пункта, позволяющий по заданному списку транспортных пунктов выполнить поиск расписаний всех рейсов, проходящих через данные транспортные пункты; результат поиска представляет собой список расписаний всех рейсов.

К группе *задач поиска и объединения пассажиров* относятся:

- (1) поиск списка предполагаемых попутчиков, позволяющий для заданного пассажира выполнить поиск сведений о поездках и для всех найденных поездок выполнить поиск пассажиров конкретного рейса; результирующий список можно ограничивать только пассажирами, занявшими соседние места, либо конкретными операциями (бронирование, оплата, регистрация и т. п.);
- (2) поиск пассажиров по пунктам назначения, позволяющий указать атрибуты рейсов либо пунктов отправления и прибытия и выполнить поиск пассажиров, вошедших либо вышедших в указанном транспортном пункте (на станции, в порту и т. д.) согласно проездным документам;
- (3) поиск пассажиров, купивших билеты совместно, позволяющий для указанного пассажира определить список пассажиров, купивших билеты либо в конкретном кассовом терминале (у конкретного кассира), либо по выбранному номеру операции бронирования;
- (4) поиск и сравнение списков пассажиров, позволяющий для заданного оператором списка пассажиров выполнить поиск пассажиров, участвовавших в нескольких поездках совместно с другими пассажирами из списка и подсчитать число «совместных» поездок;
- (5) поиск транспортных связей, позволяющий для заданного оператором списка пассажиров выполнить построение матрицы парных «участий» (под «участиями» понимаются: (а) поездка на одном рейсе с совпадением пункта

отправления и пункта прибытия; (б) поездка на одном рейсе с совпадением пункта отправления или пункта прибытия; (в) поездка на одном рейсе); матрица «участий» представляет собой таблицу, каждая ячейка которой с номером строки i и номером столбца j содержит число совместных «участий» i -го и j -го пассажиров; ячейки на диагонали содержат числа, показывающие общее число «участий» пассажира в поездках;

- (6) поиск пассажиров с устойчивым заданным маршрутом, позволяющий по заранее заданному списку пар «пункт отправления – пункт прибытия» и диапазону дат для каждой пары выполнить поиск пассажиров, выполнявших поездки с заданными характеристиками.

К группе *специализированных запросов по авто- и авиаперевозкам* относятся:

- (1) поиск пассажиров, купивших/сдавших авиабилеты совместно, позволяющий для заданного оператором пассажира выполнить поиск всех пассажиров, участвовавших в операциях (тип операции может уточняться), имеющих тот же код бронирования, что и у заданного пассажира;
- (2) поиск пассажиров, не прошедших регистрацию на авиарейс, позволяющий для заданной оператором группы рейсов выполнить поиск пассажиров, имеющих билеты на рейс, но не прошедших регистрацию;
- (3) поиск рейсов по регистрационному знаку транспортного средства, позволяющий для заданного государственного номерного знака выполнить поиск всех рейсов, выполненных транспортным средством с заданным номером; поиск может быть выполнен с нестрогим совпадением строк (по неточно заданному номеру), а также с учетом перехода латиница/кириллица в написании букв регистрационного знака. Результатом поиска является список расписаний, удовлетворяющих заданным условиям поиска.

К группе запросов *поиска артефактов, несоответствий, различий в данных о поездках* относятся:

- (1) поиск несоответствий при совпадении ФИО и номера документа, позволяющий выполнить поиск пассажиров, у которых в нескольких поездках совпадают данные фамильно-именной группы и номер документа, но отличаются другие персональные параметры или данные документа;
- (2) поиск несоответствий при совпадении ФИО и даты рождения, позволяющий выполнить поиск пассажиров, у которых в нескольких поездках совпадают данные фамильно-именной группы и дата рождения, но отличаются другие персональные параметры или данные документа.

В данной группе запросов имеются всего два вида поиска, однако ее расширение представляется возможным и целесообразным, в первую очередь за счет формализации опыта персонала (следователей, оперативных работников). В перспективе это может стать основой для создания ряда новых и перспективных технических решений.

4 Заключение

Данная статья продолжила цикл исследований, начатый в [1]. При этом если в статье [1], представляя онтологию целевой предметной области — информационного обеспечения безопасности транспортных пассажирских перевозок, авторы ограничились информационной моделью данной предметной области, то здесь эта модель уже наполнена реальными и существенными методами.

Наиболее важным здесь было предложить не просто формальные инструменты для оперирования объектами предметной области, но и сделать их максимально специализированными, т. е. пригодными и полезными для реальных задач выбранной предметной области. При этом вполне осмысленно применялось понятие «аналитический», поскольку работа потенциальных пользователей в этой области — следователей и оперативных работников — это деятельность аналитическая, хотя и со своей спецификой, а значит, здесь можно и нужно следовать порядку, принятым в аналитических системах, системах поддержки принятия решений.

Именно этот шаг и реализован в статье описанием инструментов первого этапа работы аналитика — отбора данных. При этом имеется важная иллюстрация того, что даже простые поисковые средства за счет глубины проработки предметной области позволяют решать узкоспециализированные задачи, уже претендующие на поддержку аналитической деятельности. Примерами таких средств служат запросы на поиск артефактов, которые, естественно, могут использоваться в качестве шаблонов для расширения аналитического функционала описываемой СППР. Но все же описанные инструменты — это прежде всего средства отбора данных для более глубокой обработки.

Соответствующие задачи и алгоритмы их решения названы авторами расчетными, им будет посвящена следующая статья по данной проблематике.

Литература

1. *Борисов А. В., Босов А. В., Жуков Д. В., Иванов А. В., Сушко Д. В.* Информационные аспекты обеспечения безопасности на транспорте: онтология предметной области, модели и варианты использования // Системы и средства информатики, 2020. Т. 30. № 1. С. 126–134.
2. *Каньковски П.* «Как ваша фамилия?», или русский MetaPhone // Программист, 2002. № 8. С. 36–39.
3. *Philips L.* Hanging on the Metaphone // Comput. Lang., 1990. Vol. 7. Iss. 12. P. 39–43.
4. *Navarro G.* A guided tour to approximate string matching // ACM Comput. Surv., 2001. Vol. 33. Iss. 1. P. 31–88.
5. *Левенштейн В. И.* Двоичные коды с исправлением выпадений, вставок и замещений символов // Докл. Акад. наук СССР, 1965. Т. 163. № 4. С. 845–848.

6. Попов И. А., Персианинов Ю. Н., Миронов Я. А., Дегтярев А. Ю. Использование возможностей ЕИТКС ОВД в деятельности органов предварительного следствия в системе МВД России. — М.: Проспект, 2014. 112 с.

Поступила в редакцию 09.01.21

INFORMATION ASPECTS OF SECURITY IN TRANSPORT: SEARCH AND SELECTION OF INFORMATION

A. V. Borisov, A. V. Bosov, D. V. Zhukov, and A. V. Ivanov

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

Abstract: This is the second article of a series devoted to the description of the decision support system in the field of ensuring the safety of passenger traffic and countering illegal activities in transport. Based on the conceptual model proposed in the first article, the functional content of the system is formed. A classification of analytical functions is proposed, including search queries, analytical calculations, and data processing tasks, and the content of tasks of the first (search) type specialized within the field of study is presented. Some of the general purpose search queries are combined into a group of utility methods. The second part consists of specialized transport search queries. Subject matter related searches announce the planned presentation of advanced analytic functions.

Keywords: transport safety; decision support system; search query; calculated task; data processing

DOI: 10.14357/08696527210208

References

1. Borisov, A. V., A. V. Bosov, D. V. Zhukov, A. V. Ivanov, and D. V. Sushko. 2020. Informatsionnye aspekty obespecheniya bezopasnosti na transporte: ontologiya predmetnoy oblasti, modeli i varianty ispol'zovaniya [Information aspects of security in transport: Ontology of the subject area, models and cases]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(1):126–134.
2. Kankovski, P. 2002. “Kak vasha familiya?”, ili russkiy MetaPhone [“What is your surname?”, or Russian MetaPhone]. *Programmist* [Programmer] 8:36–39.
3. Philips, L. 1990. Hanging on the Metaphone. *Comput. Lang.* 7(12):39–43.
4. Navarro, G. 2001. A guided tour to approximate string matching. *ACM Comput. Surv.* 33(1):31–88.
5. Levenshtein, V. I. 1965. Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics Doklady* 10(8):707–710.

6. Popov, I. A., Yu. N. Persianinov, Ya. A. Mironov, and A. Yu. Degtyarev. 2014. *Ispol'zovanie vozmozhnostey EITKS OVD v deyatel'nosti organov predvaritel'nogo sledstviya v sisteme MVD Rossii* [Using the capabilities of EITKS IAB in the activities of the preliminary investigation bodies in the system of the Ministry of Internal Affairs of Russia]. Moscow: Prospekt. 112 p.

Received January 9, 2021

Contributors

Borisov Andrey V. (b. 1965) — Doctor of Science in physics and mathematics, 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; ABorosov@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

Zhukov Denis V. (b. 1979) — principal specialist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; DZhukov@ipiran.ru

Ivanov Alexey V. (b. 1976) — 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; AIvanov@ipiran.ru

СТРАТЕГИЯ ИССЛЕДОВАНИЙ И РАЗРАБОТОК В ОБЛАСТИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА II: СРАВНИТЕЛЬНЫЙ АНАЛИЗ НАУКОМЕТРИЧЕСКИХ ПОКАЗАТЕЛЕЙ В МИРЕ И В РОССИЙСКОЙ ФЕДЕРАЦИИ

А. В. Борисов¹, А. В. Босов², Д. В. Жуков³

Аннотация: Статья продолжает цикл работ, открытый в предыдущем номере журнала представлением результатов анализа влияния государственного управления на эффективность проведения исследований и разработок в области искусственного интеллекта (AI R&D). Вторая часть цикла посвящена анализу значений наукометрических показателей AI R&D в мире и в Российской Федерации в период с 2000 по 2019 гг. Представлена информация о мировых индексируемых научных изданиях в области AI R&D. Даны сведения о мировой публикационной активности в области AI R&D. Обсуждаются результаты сравнительного статистического анализа наукометрических показателей исследователей из Российской Федерации вообще и Российской академии наук (РАН) в частности.

Ключевые слова: искусственный интеллект; наукометрические индексы; реферативные базы данных; платформа Scopus; платформа Web of Science

DOI: 10.14357/08696527210209

1 Введение

Первая часть [1] данного цикла статей содержала краткий исторический очерк государственного управления исследованиями и разработками в области искусственного интеллекта в СССР и Российской Федерации, которые можно считать некоторой аппаратно-технологической базой реализации действующей в РФ «Национальной стратегии развития искусственного интеллекта на период до 2030 года» [2] (далее — Стратегия).

Несмотря на то что основное внимание в Стратегии уделяется технологическому и внедренческому аспектам, залог ее успешного выполнения кроется в возможности РФ вести самостоятельные фундаментальные исследования по основным направлениям AI. Для объективности оценки научного потенциала

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

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

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

предлагается рассмотреть наукометрические показатели РФ в области AI R&D в международных реферативных базах данных публикаций и патентов.

Проведенное статистическое исследование затрагивает временной отрезок с 2000 по 2019 гг. Выбор верхней границы обусловлен существенными (год и более) задержками в индексировании публикаций в реферативных базах, так что на текущий момент достоверными можно считать только сведения по 2019 г. включительно. Кроме того, поскольку источники реферативной информации постоянно обновляются, использованные авторами конкретные статистические данные были актуальными лишь на день выполнения запросов. На данный момент величины наукометрических показателей изменились, а в использованные значения, таким образом, включены погрешности, связанные, например, с частичным учетом публикаций 2020 года при неполной статистике по 2019 году. Представляется, что такого рода издержки совершенно несущественны в рамках данной работы, а сделанные выводы и выявленные тенденции верны и актуальны и в настоящее время.

Результаты исследований изложены в статье следующим образом. В разд. 2 представлена информация о мировых индексируемых научных изданиях в области AI R&D. Сведения о мировой публикационной активности AI R&D даны в разд. 3. Разделы 4 и 5 посвящены соответственно статистическому анализу наукометрических показателей исследователей из Российской Федерации вообще и Российской академии наук в частности. Выводы по выполненному анализу представлены в заключении (разд. 6).

2 Мировые академические издания в области AI R&D

Согласно поисковой интернет-платформе Scopus [3], объединяющей реферативные базы данных публикаций в научных журналах и патентов, на начало 2020 г. в мире насчитывалось 173 значимых научных журнала, публикующих статьи по тематике AI R&D (табл. 1). Из них только два журнала: «Труды СПИИРАН» (SPIIRAS Proceedings, Q3) и «Cybernetics and Physics» (англоязычный журнал, Q4) издаются в России.

Из табл. 1 видно, что Россия по числу специализированных научных журналов занимает 13-е место с менее чем 1,2% журналов от общемирового объема.

3 Анализ международной публикационной активности

Статистический анализ публикаций по тематике AI R&D выполнялся на основе данных поисковой интернет-платформы Web of Science (WoS) [4], объединяющей реферативные базы данных публикаций в научных журналах и патентов.

За период 2000–2019 гг. был выбран корпус публикаций, соответствующих следующим ограничениям (перечислены ключевые слова):

Таблица 1 Распределение журналов AI R&D по государствам

Место	Государство	Число журналов AI R&D
1	Нидерланды	43
2	США	37
3	Великобритания	29
4	ФРГ	12
5	КНР	7
6	Швейцария	6
7	Сингапур	6
8	Южная Корея	4
9	Индия	4
10	Тайвань	3
11	Польша	3
12	Япония	3
13	Россия	2
14	Иран	2
15	Чехия	2

(*ts = “artificial intelligence” or ts = “deep learning” or ts = “machine learning” or ts = “neural networks” or ts = “computer vision” or ts = “robotics” or ti = “artificial intelligence” or ti = “deep learning” or ti = “machine learning” or ti = “neural networks” or ti = “computer vision” or ti = “robotics”*)

При этом из запроса были исключены следующие категории:

Kamezopuu web of science: (physics condensed matter or toxicology or genetics heredity or optics or information science library science or engineering chemical or medicine research experimental or engineering marine or biotechnology applied microbiology or mining mineral processing or materials science multidisciplinary or oncology or mechanics or physics particles fields or agriculture multidisciplinary or physics applied or biophysics or instruments instrumentation or chemistry analytical or biology or engineering biomedical or water resources or engineering geological or physics multidisciplinary or chemistry medicinal or forestry or oceanography or geology or physiology or microbiology or spectroscopy or physics nuclear or physics mathematical or geochemistry geophysics or agronomy or biochemical research methods or nuclear science technology or biochemistry molecular biology or clinical neurology or energy fuels or critical care medicine or meteorology atmospheric sciences or geography physical or ecology or robotics or pharmacology pharmacy or chemistry multidisciplinary or radiology nuclear medicine medical imaging or environmental sciences or chemistry organic or chemistry physical or astronomy astrophysics or limnology or geriatrics gerontology or history philosophy of science or materials science characterization testing or physics fluids plasmas or international relations or philosophy or cell biology or criminology penology or engineering ocean or area studies or health care sciences services or engineering mechanical or medicine

Таблица 2 Распределение публикаций AI R&D по государствам

Место	Государство	Число публикаций	Доля от общего числа, %
1	КНР	59 855	22,634
2	США	55 759	21,085
3	Великобритания	14 820	5,604
4	Индия	14 799	5,596
5	ФРГ	11 936	4,513
6	Испания	10 717	4,053
7	Япония	10 643	4,025
8	Канада	10 306	3,897
9	Италия	9 450	3,573
10	Франция	8 820	3,335
11	Австралия	8 186	3,095
12	Южная Корея	7 697	2,911
14	Тайвань	5 813	2,198
15	Турция	5 737	2,169
16	Иран	4 980	1,883
17	Польша	4 698	1,777
18	Сингапур	4 114	1,556
19	Швейцария	3 408	1,289
20	Греция	3 257	1,232
21	Нидерланды	3 235	1,223
22	Россия	3 146	1,19
23	Мексика	3 065	1,159
24	Малайзия	2 960	1,119
25	Португалия	2 613	0,988

Примечания.

- 162 значения показателя «Страны/регионы» вышло за пределы параметров отображения.
- 3 283 записи (1,241%) не содержат данных в анализируемом поле.

legal or ophthalmology or polymer science or geosciences multidisciplinary or public administration or zoology or biodiversity conservation)

Параметры запросов к системе приведены в оригинальном исходном виде, чтобы обеспечить их воспроизводимость. Первая часть запроса определяет набор тем и словосочетаний в названиях публикаций: artificial intelligence (искусственный интеллект), deep learning (глубокое обучение), machine learning (машинное обучение), neural networks (нейронные сети), computer vision (компьютерное зрение), robotics (робототехника). Второй блок содержит перечень тем-исключений, не относящихся непосредственно к обсуждаемой.

Результатом данного запроса стали около 260 000 зарегистрированных публикаций. Распределение общемировых публикаций по государствам представлено в табл. 2, по годам — в табл. 3.

Распределение публикаций по организациям, отраслям и источникам финансирования приведено в табл. 4, 5 и 6 соответственно.

Таблица 3 Распределение общемировых публикаций AI R&D по годам

Год издания	Число публикаций	Доля от общего числа, %
2019	36 727	13,888
2018	37 447	14,16
2017	28 931	10,94
2016	21 779	8,236
2015	17 522	6,626
2014	12 908	4,881
2013	10 954	4,142
2012	9 361	3,54
2011	8 777	3,319
2010	8 448	3,195
2009	10 205	3,859
2008	8 992	3,4
2007	8 165	3,088
2006	7 828	2,96
2005	6 623	2,504
2004	5 707	2,158
2003	5 167	1,954
2002	4 606	1,742
2001	4 061	1,536
2000	4 338	1,64

Таблица 4 Распределение мировых публикаций AI R&D по организациям

Место	Организация	Государство	Число публикаций	Доля от общего числа, %
1	CHINESE ACADEMY OF SCIENCES	КНР	5 834	2,206
2	UNIVERSITY OF CALIFORNIA SYSTEM	США	4 423	1,673
3	CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE CNRS	Франция	3 348	1,266
4	TSINGHUA UNIVERSITY	КНР	2 454	0,928
5	INDIAN INSTITUTE OF TECHNOLOGY SYSTEM IIT SYSTEM	Индия	2 187	0,827
6	STATE UNIVERSITY SYSTEM OF FLORIDA	США	2 150	0,813
7	UNIVERSITY OF LONDON	Великобритания	2 037	0,77
8	UNIVERSITY OF TEXAS SYSTEM	США	2 014	0,762
9, 10	NANYANG TECHNOLOGICAL UNIVERSITY, NATIONAL INSTITUTE OF EDUCATION NIE SINGAPORE	Сингапур	1 961	0,742
11	UNIVERSITY SYSTEM OF GEORGIA	США	1 952	0,738

Окончание табл. 4 на с. 94

Таблица 4 (окончание) Распределение мировых публикаций AI R&D по организациям

Место	Организация	Государство	Число публикаций	Доля от общего числа, %
12	CARNEGIE MELLON UNIVERSITY	США	1 772	0,67
13, 14	SHANGHAI JIAO TONG UNIVERSITY, SOUTHEAST UNIVERSITY CHINA	КНР	1 718	0,65
15	HARBIN INSTITUTE OF TECHNOLOGY	КНР	1 666	0,65
16	UNIVERSITY OF CHINESE ACADEMY OF SCIENCES CAS	КНР	1 657	0,627
17	MASSACHUSETTS INSTITUTE OF TECHNOLOGY MIT	США	1 648	0,623
18	UNIVERSITY OF ELECTRONIC SCIENCE TECHNOLOGY OF CHINA	КНР	1 545	0,584
19	PENNSYLVANIA COMMONWEALTH SYSTEM OF HIGHER EDUCATION	США	1 504	0,569
20	ZHEJIANG UNIVERSITY	КНР	1 419	0,537
21	MICROSOFT	США	1 415	0,535
22	HUAZHONG UNIVERSITY OF SCIENCE TECHNOLOGY	КНР	1 404	0,531
23	NATIONAL UNIVERSITY OF SINGAPORE	Сингапур	1 402	0,65
24	BEIHANG UNIVERSITY	КНР	1 390	0,526
25	STANFORD UNIVERSITY	США	1 373	0,519

Примечания.

1. 45 774 значения показателя «Профили организаций» вышло за пределы параметров отображения.
2. 3 436 записей (1,299%) не содержат данных в анализируемом поле.

Таблица 5 Распределение мировых публикаций AI R&D по отраслям

Место	Категория	Число публикаций	Доля от общего числа, %
1	Вычислительная техника и AI	98 557	37,268
2	Электроника и электротехника	84 383	31,909
3	Теоретические методы вычислительной техники	52 394	19,812
4	Информационные системы вычислительной техники	36 993	13,989
5	Вычислительная техника: междисциплинарные статьи	21 923	8,29
6	Системы автоматического управления	20 939	7,918
7	Телекоммуникации	19 429	7,347

Окончание табл. 5 на с. 95

Таблица 5 (окончание) Распределение мировых публикаций AI R&D по отраслям

Место	Категория	Число публикаций	Доля от общего числа, %
8	Вычислительная техника: разработка программного обеспечения	15 437	5,837
9	Вычислительная техника: аппаратная архитектура	12 510	4,731
10	Робототехника	11 414	4,316
11	Обработка изображений, технологии фотографии	11 162	4,221
12	Нейронауки	8 189	3,097
13	Инженерные науки (междисциплинарные)	7 765	2,936
14	Вычислительная техника: кибернетика	7 653	2,894
15	Оптика	7 369	2,787
16	Исследование операций	7 132	2,697
17	Прикладная математика	7 089	2,681
18	Математическая вычислительная биология	6 487	2,453
19	Биомедицинская инженерия	6 276	2,373
20	Приборостроение	4 483	1,695
21	Дистанционное зондирование	3 673	1,389
22	Биохимические методы исследования	3 588	1,357
23	Индустриальная инженерия	3 521	1,331
24	Акустика	3 346	1,265
25	Медицинская информатика	3 201	1,21

Примечания.

- 175 значений показателя «Категории Web of Science» вышло за пределы параметров отображения.
- 601 запись (0,227%) не содержит данных в анализируемом поле.

Таблица 6 Распределение публикаций по источникам финансирования

Место	Финансирующие организации	Государство	Число публикаций	Доля от общего числа, %
1	NATIONAL NATURAL SCIENCE FOUNDATION OF CHINA	КНР	28 114	10,631
2	NATIONAL SCIENCE FOUNDATION NSF	США	7 974	3,015
3	UNITED STATES DEPARTMENT OF HEALTH HUMAN SERVICES	США	4 705	1,779
4	NATIONAL INSTITUTES OF HEALTH NIH USA	США	4 658	1,761
5	FUNDAMENTAL RESEARCH FUNDS FOR THE CENTRAL UNIVERSITIES	КНР	4 054	1,533
6	EUROPEAN UNION EU	ЕС	3 527	1,334
7	NATIONAL BASIC RESEARCH PROGRAM OF CHINA	КНР	2 254	0,852

Окончание табл. 6 на с. 96

Таблица 6 (окончание) Распределение публикаций по источникам финансирования

Место	Финансирующие организации	Государство	Число публикаций	Доля от общего числа, %
8	MINISTRY OF EDUCATION CULTURE SPORTS SCIENCE AND TECHNOLOGY JAPAN MEXT	Япония	2 035	0,77
9	UNITED STATES DEPARTMENT OF DEFENSE	США	2 011	0,76
10	NATURAL SCIENCES AND ENGINEERING RESEARCH COUNCIL OF CANADA	Канада	1 959	0,741
11	CHINA POSTDOCTORAL SCIENCE FOUNDATION	КНР	1 800	0,681
12	ENGINEERING PHYSICAL SCIENCES RESEARCH COUNCIL EPSRC	Великобритания	1 767	0,668
13	JAPAN SOCIETY FOR THE PROMOTION OF SCIENCE	Япония	1 612	0,61
14	NATIONAL COUNCIL FOR SCIENTIFIC AND TECHNOLOGICAL DEVELOPMENT CNPQ	Бразилия	1 550	0,586
15	NATIONAL KEY RESEARCH AND DEVELOPMENT PROGRAM OF CHINA	КНР	1 505	0,569
16	GERMAN RESEARCH FOUNDATION DFG	ФРГ	1 414	0,535
17	SPANISH GOVERNMENT	Испания	1 375	0,52
18	AUSTRALIAN RESEARCH COUNCIL	Австралия	1 271	0,481
19	NATURAL SCIENCE FOUNDATION OF JIANGSU PROVINCE	КНР	1 166	0,441
20	OFFICE OF NAVAL RESEARCH	США	1 136	0,43
21	GRANTS IN AID FOR SCIENTIFIC RESEARCH KAKENHI	Япония	1 081	0,409
22	CAPEX	Франция	1 057	0,4
23	NATIONAL HIGH TECHNOLOGY RESEARCH AND DEVELOPMENT PROGRAM OF CHINA	КНР	982	0,371
24	NATIONAL KEY R D PROGRAM OF CHINA	КНР	950	0,359
25	CHINA SCHOLARSHIP COUNCIL	КНР	917	0,347

Примечания.

1. 70 777 значений показателя «Финансирующие организации» вышло за пределы параметров отображения.
2. 162 130 записей (61,308%) не содержат данных в анализируемом поле.

Из приведенных выше статистических данных видно, что на начало 2020 г. ни одна из российских организаций/объединений/корпораций (академических, образовательных или производственных/предпринимательских) не входит в число значимых центров исследований в области AI R&D и не представлена в качестве источника их финансирования.

4 Анализ российской публикационной активности

Для получения данных по российским публикациям в области AI R&D, зарегистрированным в международной базе WoS, первая часть поискового запроса была скорректирована дополнительными условиями о государственной принадлежности авторов: (cu = "Russia" or cu = "Russian Federation").

Результатом данного запроса стали более 3 100 зарегистрированных публикаций. Распределение российских публикаций по годам, организациям и отраслям представлено в табл. 7, 8 и 9 соответственно.

Таблица 7 Распределение российских публикаций AI R&D по годам

Год издания	Число публикаций	Доля от общего числа, %
2019	572	18,182
2018	691	21,964
2017	551	17,514
2016	373	11,856
2015	204	6,484
2014	116	3,687
2013	64	2,034
2012	46	1,462
2011	48	1,526
2010	37	1,176
2009	38	1,208
2008	49	1,558
2007	46	1,462
2006	29	0,922
2005	43	1,367
2004	40	1,271
2003	33	1,049
2002	38	1,208
2001	38	1,208
2000	31	0,985

Таблица 8 Распределение российских публикаций AI R&D по организациям

№	Organization profile (original)	Профиль организации	Число публикаций	Доля от общего числа, %
1	RUSSIAN ACADEMY OF SCIENCES	РАН	761	24,189
2	ITMO UNIVERSITY	ИТМО	230	7,311
3	NATIONAL RESEARCH UNIVERSITY HIGHER SCHOOL OF ECONOMICS	НИУ ВШЭ	185	5,88

Продолжение табл. 8 на с. 98

Таблица 8 (продолжение) Распределение российских публикаций AI R&D по организациям

№	Organization profile (original)	Профиль организации	Число публикаций	Доля от общего числа, %
4	LOMONOSOV MOSCOW STATE UNIVERSITY	МГУ	181	5,753
5	MOSCOW INSTITUTE OF PHYSICS TECHNOLOGY	МФТИ	122	3,878
6	PETER THE GREAT ST PETERSBURG POLYTECHNIC UNIVERSITY	СПБПТУ	100	3,179
7	SAINT PETERSBURG STATE UNIVERSITY	СПБУ	94	2,988
8	SKOLKOVO INSTITUTE OF SCIENCE TECHNOLOGY	Сколтех	86	2,734
9	URAL FEDERAL UNIVERSITY	УФУ	86	2,734
10	KAZAN FEDERAL UNIVERSITY	КазФУ	83	2,638
11	SAINT PETERSBURG STATE ELECTROTECHNICAL UNIVERSITY	ЛЭТИ	82	2,606
12	FEDERAL RESEARCH CENTER COMPUTER SCIENCE CONTROL OF RAS	ФИЦ ИУ РАН	79	2,511
13	NATIONAL RESEARCH NUCLEAR UNIVERSITY MEPhI MOSCOW ENGINEERING PHYSICS INSTITUTE	МИФИ	78	2,479
14	ST PETERSBURG INSTITUTE OF INFORMATICS AUTOMATION	СПИИ РАН	71	2,257
15	LOBACHEVSKY STATE UNIVERSITY OF NIZHNI NOVGOROD	НИНГУ	61	1,939
16	NATIONAL RESEARCH CENTRE KURCHATOV INSTITUTE	НИЦ «Курчатовский институт»	59	1,875
17	TOMSK POLYTECHNIC UNIVERSITY	ТПУ	55	1,748
18	KHARKEVICH INSTITUTE FOR INFORMATION TRANSMISSION PROBLEMS OF THE RAS	ИППИ РАН	54	1,716
19	V A TRAPEZNIKOV INSTITUTE OF CONTROL SCIENCES RUSSIAN ACADEMY OF SCIENCES	ИПУ РАН	52	1,653
20	BAUMAN MOSCOW STATE TECHNICAL UNIVERSITY	МГТУ	43	1,367
21	SOUTHERN FEDERAL UNIVERSITY	ЮжФУ	42	1,335
22	NOVOSIBIRSK STATE UNIVERSITY	НГУ	41	1,303
23	NOVOSIBIRSK STATE TECHNICAL UNIVERSITY	НГПУ	40	1,271
24	TOMSK STATE UNIVERSITY	ТГУ	39	1,24

Окончание табл. 8 на с. 99

Таблица 8 (окончание) Распределение российских публикаций AI R&D по организациям

№	Organization profile (original)	Профиль организации	Число публикаций	Доля от общего числа, %
25	INNOPOLIS UNIVERSITY	Иннополис (Казань)	38	1,208

Примечания.

- 2 191 значение показателя «Профили организаций» вышло за пределы параметров отображения.
- 0 записей (0,000%) не содержат данных в анализируемом поле.

Таблица 9 Распределение российских публикаций AI R&D по отраслям

№	Web of Science Category (original)	Категория	Число публикаций	Доля от общего числа, %
1	COMPUTER SCIENCE ARTIFICIAL INTELLIGENCE	Вычислительная техника и ИИ	885	28,131
2	ENGINEERING ELECTRICAL ELECTRONIC	Электроника и электротехника	681	21,647
3	COMPUTER SCIENCE THEORY METHODS	Теоретические методы вычислительной техники	612	19,453
4	COMPUTER SCIENCE INFORMATION SYSTEMS	Информационные системы вычислительной техники	345	10,966
5	OPTICS	Оптика	249	7,915
6	AUTOMATION CONTROL SYSTEMS	Системы автоматического управления	212	6,739
7	COMPUTER SCIENCE INTERDISCIPLINARY APPLICATIONS	Вычислительная техника: междисциплинарные статьи	204	6,484
8	TELECOMMUNICATIONS	Телекоммуникации	166	5,277
9	PHYSICS MULTIDISCIPLINARY	Физика: междисциплинарные статьи	133	4,228
10	MATHEMATICS APPLIED	Прикладная математика	125	3,973
11	ROBOTICS	Робототехника	121	3,846
12	COMPUTER SCIENCE SOFTWARE ENGINEERING	Вычислительная техника: разработка программного обеспечения	111	3,528
13	IMAGING SCIENCE PHOTO- GRAPHIC TECHNOLOGY	Обработка изображений, технологии фотографии	105	3,338
14	NEUROSCIENCES	Нейронауки	102	3,242
15	PHYSICS APPLIED	Прикладная физика	101	3,21

Окончание табл. 9 на с. 100

Таблица 9 (окончание) Распределение российских публикаций AI R&D по отраслям

№	Web of Science Category (original)	Категория	Число публикаций	Доля от общего числа, %
16	INSTRUMENTS INSTRUMENTATION	Приборостроение	99	3,147
17	COMPUTER SCIENCE HARDWARE ARCHITECTURE	Вычислительная техника: аппаратная архитектура	95	3,02
18	COMPUTER SCIENCE CYBERNETICS	Вычислительная техника: кибернетика	86	2,734
19	EDUCATION EDUCATIONAL RESEARCH	Образование, образова- тельные исследования	59	1,875
20	ENGINEERING MULTIDISCIPLINARY	Инженерия: междисцип- линарные статьи	59	1,875
21	OPERATIONS RESEARCH MANAGEMENT SCIENCE	Исследование операций	59	1,875
22	MATHEMATICS	Математика	56	1,78
23	ENGINEERING INDUSTRIAL	Индустриальная инженерия	52	1,653
24	MATHEMATICAL COMPUTATIONAL BIOLOGY	Математическая вычис- лительная биология	48	1,526
25	SOCIAL SCIENCES INTERDISCIPLINARY	Социальные науки: междисциплинарные статьи	43	1,367

Примечания.

- 106 значений показателя «Категории Web of Science» вышло за пределы параметров отображения.
- 31 запись (0,985%) не содержит данных в анализируемом поле.

5 Анализ публикационной активности Российской академии наук

Для получения данных по публикациям РАН в области AI R&D, зарегистрированным в международной базе WoS, первая часть поискового запроса была скорректирована дополнительными условиями об организационной принадлежности авторов публикации: (og = “Russian academy of sciences” or oo = “Russian academy of sciences”) Результатом данного запроса стали более 1000 зарегистрированных публикаций. Распределение академических публикаций по годам, источникам финансирования и отраслям представлено в табл. 10, 11 и 12 соответственно.

В табл. 13 по убыванию рейтинга актуальности представлено процентное соотношение исследований по различным отраслям AI R&D в мире, Российской Федерации и РАН.

Таблица 10 Распределение публикаций РАН в области AI R&D по годам

Год издания	Число публикаций	Доля от общего числа, %
2019	169	15,928
2018	211	19,887
2017	125	11,781
2016	91	8,577
2015	60	5,655
2014	43	4,053
2013	23	2,168
2012	23	2,168
2011	18	1,697
2010	23	2,168
2009	25	2,356
2008	29	2,733
2007	36	3,393
2006	24	2,262
2005	28	2,639
2004	21	1,979
2003	23	2,168
2002	21	1,979
2001	22	2,074
2000	22	2,074

Таблица 11 Распределение публикаций РАН по источникам финансирования

Место	Финансирующие организации	Государство	Число публикаций	Доля от общего числа, %
1	RUSSIAN FOUNDATION FOR BASIC RESEARCH RFBR	Россия	237	22,337
2	RUSSIAN SCIENCE FOUNDATION RSF	Россия	134	12,63
3	MINISTRY OF EDUCATION AND SCIENCE RUSSIAN FEDERATION	Россия	50	4,713
4	RUSSIAN ACADEMY OF SCIENCES	Россия	33	3,11
5	NATIONAL INSTITUTES OF HEALTH NIH USA	США	18	1,697
6	UNITED STATES DEPARTMENT OF HEALTH HUMAN SERVICES	США	18	1,697
7	GOVERNMENT OF THE RUSSIAN FEDERATION	Россия	15	1,414
8	EUROPEAN UNION EU	ЕС	14	1,32
9	NATIONAL SCIENCE FOUNDATION NSF	США	14	1,32
10	GERMAN RESEARCH FOUNDATION DFG	ФРГ	13	1,225

Окончание табл. 11 на с. 102

Таблица 11 (окончание) Распределение публикаций РАН по источникам финансирования

Место	Финансирующие организации	Государство	Число публикаций	Доля от общего числа, %
11	RUSSIAN FEDERATION	Россия	11	1,037
12	NATIONAL NATURAL SCIENCE FOUNDATION OF CHINA	КНР	10	0,943
13	PORTUGUESE FOUNDATION FOR SCIENCE AND TECHNOLOGY	Португалия	10	0,943
14	ENGINEERING PHYSICAL SCIENCES RESEARCH COUNCIL EPSRC	Великобритания	9	0,848
15	EUROPEAN RESEARCH COUNCIL ERC	ЕС	9	0,848
16	MINISTRY OF EDUCATION CULTURE SPORTS SCIENCE AND TECHNOLOGY JAPAN MEXT	Япония	8	0,754
17	SWISS NATIONAL SCIENCE FOUNDATION SNSF	Швейцария	8	0,754
18	ALEXANDER VON HUMBOLDT FOUNDATION	Германия	7	0,66
19	SCIENCE TECHNOLOGY FACILITIES COUNCIL STFC	Великобритания	7	0,66
20	ACADEMY OF FINLAND	Финляндия	6	0,566
21	CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE CNRS	Франция	6	0,566
22	EURATOM RESEARCH AND TRAINING PROGRAMME	ЕС	6	0,566
23	JAPAN SOCIETY FOR THE PROMOTION OF SCIENCE	Япония	6	0,566

Таблица 12 Распределение публикаций РАН в области AI R&D по отраслям

Место	Web of Science Category (original)	Категория	Число публикаций	Доля от общего числа, %
1	COMPUTER SCIENCE ARTIFICIAL INTELLIGENCE	Вычислительная техника и ИИ	192	42,105
2	COMPUTER SCIENCE THEORY METHODS	Теоретические методы вычислительной техники	113	24,781
3	ENGINEERING ELECTRICAL ELECTRONIC	Электроника и электротехника	79	17,325

Окончание табл. 12 на с. 103

Таблица 12 (окончание) Распределение публикаций РАН в области AI R&D по отраслям

Место	Web of Science Category (original)	Категория	Число публикаций	Доля от общего числа, %
4	COMPUTER SCIENCE INFORMATION SYSTEMS	Информационные системы вычислительной техники	48	10,526
5	COMPUTER SCIENCE CYBERNETICS	Вычислительная техника: кибернетика	33	7,237
6	MATHEMATICS APPLIED	Прикладная математика	30	6,579
7	TELECOMMUNICATIONS	Телекоммуникации	30	6,579
8	COMPUTER SCIENCE INTERDISCIPLINARY APPLICATIONS	Вычислительная техника: междисциплинарные статьи	28	6,14
9	AUTOMATION CONTROL SYSTEMS	Системы автоматического управления	24	5,263
10	MULTIDISCIPLINARY SCIENCES	Междисциплинарные статьи	22	4,825
11	COMPUTER SCIENCE SOFTWARE ENGINEERING	Вычислительная техника: разработка программного обеспечения	21	4,605
12	MATHEMATICS	Математика	21	4,605
13	NEUROSCIENCES	Нейронауки	20	4,386
14	COMPUTER SCIENCE HARDWARE ARCHITECTURE	Вычислительная техника: аппаратная архитектура	16	3,509
15	ENGINEERING MULTIDISCIPLINARY	Инженерия: междисциплинарные статьи	13	2,851
16	MATHEMATICAL COMPUTATIONAL BIOLOGY	Математическая вычислительная биология	12	2,632
17	OPERATIONS RESEARCH MANAGEMENT SCIENCE	Исследование операций	8	1,754
18	MATHEMATICS INTERDISCIPLINARY APPLICATIONS	Математика: междисциплинарные статьи	6	1,316
19	ECONOMICS	Экономика	5	1,096
20	IMAGING SCIENCE PHOTOGRAPHIC TECHNOLOGY	Обработка изображений, технологии фотографии	5	1,096
21	REMOTE SENSING	Удаленное зондирование	5	1,096
22	BUSINESS	Бизнес	4	0,877
23	ACOUSTICS	Акустика	3	0,658
24	STATISTICS PROBABILITY	Вероятность и статистика	3	0,658
25	BEHAVIORAL SCIENCES	Бихевиористика	2	0,439

Примечания.

- 27 значений показателя «Категории Web of Science» вышло за пределы параметров отображения.
- 12 записей (2,632%) не содержат данных в анализируемом поле.

Таблица 13 Распределение приоритетов исследований в мире, РФ и РАН

Приоритет	Web of Science Category (original)	Категория	Доля, %		
			в мире	в РФ	в РАН
1	COMPUTER SCIENCE ARTIFICIAL INTELLIGENCE	Вычислительная техника и ИИ	37,268	28,131	42,105
2	ENGINEERING ELECTRICAL ELECTRONIC	Электроника и электротехника	31,909	21,647	17,325
3	COMPUTER SCIENCE THEORY METHODS	Теоретические методы вычислительной техники	19,812	19,453	24,781
4	COMPUTER SCIENCE INFORMATION SYSTEMS	Информационные системы вычислительной техники	13,989	10,966	10,526
5	COMPUTER SCIENCE INTERDISCIPLINARY APPLICATIONS	Вычислительная техника: междисциплинарные статьи	8,29	6,484	6,14
6	AUTOMATION CONTROL SYSTEMS	Системы автоматического управления	7,918	6,739	5,263
7	TELECOMMUNICATIONS	Телекоммуникации	7,347	5,277	6,579
8	COMPUTER SCIENCE SOFTWARE ENGINEERING	Вычислительная техника: разработка программного обеспечения	5,837	3,528	4,605
9	COMPUTER SCIENCE HARDWARE ARCHITECTURE	Вычислительная техника: аппаратная архитектура	4,731	3,02	3,509
11	IMAGING SCIENCE PHOTOGRAPHIC TECHNOLOGY	Обработка изображений, технологии фотографии	4,221	3,338	1,096
12	NEUROSCIENCES	Нейронауки	3,097	3,242	4,386
13	ENGINEERING MULTIDISCIPLINARY	Инженерия: междисциплинарные статьи	2,936	1,875	2,851
14	COMPUTER SCIENCE CYBERNETICS	Вычислительная техника: кибернетика	2,894	2,734	7,237
16	OPERATIONS RESEARCH MANAGEMENT SCIENCE	Исследование операций	2,697	1,875	1,754
17	MATHEMATICS APPLIED	Прикладная математика	2,681	3,973	6,579
18	MATHEMATICAL COMPUTATIONAL BIOLOGY	Математическая вычислительная биология	2,453	1,526	2,632

6 Заключение

По результатам проведенного статистического анализа наукометрических показателей в области AI R&D можно сделать следующие выводы.

1. Российская Федерация занимает 22-е место по публикационной активности в области AI R&D за последние 20 лет — 1,19% от общего числа публикаций.
2. По числу публикаций ни одна из российских организаций не входит в число 25 крупнейших организаций-источников публикаций в области AI R&D.
3. По числу упоминаний источников финансирования исследований ни одна из российских организаций не входит в число 25 крупнейших спонсоров исследований в области AI R&D.
4. Доля статей, опубликованных представителями институтов РАН, составляет более 34% российских публикаций.
5. Среди 25 организаций, спонсировавших наибольшее число статей по тематике AI R&D, нет ни одной коммерческой организации — только академические и образовательные.
6. Согласно информации, указанной в статьях, опубликованных в 2000–2019 гг., не более 47% публикаций РАН в области AI R&D финансировались российскими государственными организациями. Остальные источники финансирования — зарубежные.
7. Анализируя распределение публикаций по отраслям AI R&D, можно сделать вывод, что 18 из 25 наиболее актуальных мировых отраслей также входят в круг интересов российских ученых, в том числе сотрудников РАН.
8. Для практических разработок в области AI R&D в интересах органов государственной власти, в особенности в областях обороны и безопасности, вполне возможно привлечение российских ученых — сотрудников академических и образовательных организаций. Рост числа их публикаций выше общемирового, а сфера интересов вполне покрывает современные направления развития AI R&D.

Литература

1. Борисов А. В., Босов А. В., Жуков Д. В. Стратегия исследований и разработок в области искусственного интеллекта I: основные понятия и краткая хронология // Системы и средства информатики, 2021. Т. 31. № 1. С. 57–68.
2. Национальная стратегия развития искусственного интеллекта на период до 2030 года // О развитии искусственного интеллекта в Российской Федерации: Указ Президента РФ от 10 октября 2019 г. № 490. <http://static.kremlin.ru/media/events/files/ru/AN4x6HgKWANwVtMOfPDhcbRpvd1HCCsv.pdf>
3. Scopus. <https://www.scopus.com>.
4. Web of Science. <https://www.webofknowledge.com>.

Поступила в редакцию 13.01.21

**RESEARCH AND DEVELOPMENT STRATEGY IN THE FIELD
OF ARTIFICIAL INTELLIGENCE II: COMPARATIVE ANALYSIS
OF SCIENTOMETRIC INDICATORS IN THE WORLD
AND IN THE RUSSIAN FEDERATION**

A. V. Borisov, A. V. Bosov, and D. V. Zhukov

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

Abstract: The paper continues the series of works initiated in the previous issue by the presentation of the results of the analysis of the impact of public administration on the effectiveness of research and development in the field of artificial intelligence (AI R&D). This second part of the cycle is devoted to the analysis of AI R&D scientometric indicators carried out in the world and in the Russian Federation in the period from 2000 to 2019. Information on the world indexed scientific publications in the field of AI R&D is presented. Information about the world publication activity in the field of AI R&D is provided. The results of a comparative statistical analysis of scientometric indicators of researchers from the Russian Federation and, specifically, employees of the Russian Academy of Sciences, are discussed.

Keywords: artificial intelligence; scientometric indices; abstract databases; Scopus platform; Web of Science platform

DOI: 10.14357/08696527210209

References

1. Borisov, A.V., A.V. Bosov, and D.V. Zhukov. 2021. Strategiya issledovaniy i razrabotok v oblasti iskusstvennogo intellekta I: osnovnye ponyatiya i kratkaya khronologiya [Research and development strategy in the field of artificial intelligence I: Basic concepts and brief chronology]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 31(1):57–68.
2. Natsional'naya strategiya razvitiya iskusstvennogo intellekta na period do 2030 goda [National strategy for the development of artificial intelligence for the period up to 2030]. 10.10.2019. O razvitii iskusstvennogo intellekta v Rossiyskoy Federatsii: ukaz Prezidenta No. 490 [About strategy of scientific and technological development of the Russian Federation. Presidential Decree No. 490]. Available at: <http://static.kremlin.ru/media/events/files/ru/AH4x6HgKWANwVtMOfPDhcbRpvd1HCCsv.pdf> (accessed March 25, 2021).
3. Scopus. Available at: <https://www.scopus.com> (accessed March 25, 2021).
4. Web of Science. Available at: <https://www.webofknowledge.com> (accessed March 25, 2021).

Received January 13, 2021

Contributors

Borisov Andrey V. (b. 1965) — Doctor of Science in physics and mathematics, 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; ABorosov@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

Zhukov Denis V. (b. 1979) — principal specialist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; DZhukov@ipiran.ru

АНАЛИТИЧЕСКИЕ СВОЙСТВА И АСПЕКТЫ ВЫЧИСЛЕНИЯ ГАММА-ЭКСПОНЕНЦИАЛЬНОЙ ФУНКЦИИ*

М. О. Воронцов¹, А. А. Кудрявцев², С. Я. Шоргин³

Аннотация: Смеси вероятностных распределений играют важную роль в современном анализе и моделировании сложных процессов. Большое внимание исследователей традиционно уделяется распределениям, относящимся к гамма-классу. Основные вероятностные характеристики масштабных смесей обобщенных гамма-распределений не выражаются в элементарных функциях, что усложняет процесс аналитического исследования и зачастую приводит к неоправданно большим вычислительным сложностям. В работе анализируются свойства и аспекты вычисления специальной гамма-экспоненциальной функции, зарекомендовавшей себя в качестве удобного инструмента при исследовании смесей обобщенных гамма-распределений. Приводимые результаты существенно расширяют возможности применения функций типа Миттаг-Леффлера для анализа вероятностных распределений.

Ключевые слова: гамма-экспоненциальная функция; обобщенное гамма-распределение; смешанные распределения; вычислительные алгоритмы

DOI: 10.14357/08696527210210

1 Введение

В настоящее время важную роль в изучении реальных процессов и явлений играют масштабные смеси распределений.

Так, в работе [1] рассматривается масштабная смесь распределения Эрланга и распределения, плотность которого обладает свойством гиперболической монотонности. Показано, что данная смесь имеет распределение из класса обобщенных гамма-сверток, а полученные результаты могут найти широкое применение при исследовании обобщенных пуассоновских сумм и процессов Леви. В статье [2] найдены новые способы представления обобщенного распределения Линника в виде масштабной смеси распределений. В частности, показано,

* Работа выполнена при частичной финансовой поддержке РФФИ (проект 20-07-00655); исследования проводились в рамках программы Московского центра фундаментальной и прикладной математики.

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

²Московский государственный университет им. М. В. Ломоносова, факультет вычислительной математики и кибернетики, nubigena@mail.ru

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

что обобщенное распределение Линника представимо в виде смеси нормального распределения с обобщенным распределением Миттаг-Леффлера. Указанные представления применяются для доказательства ряда предельных теорем, в которых распределение Линника играет роль предельного распределения. В работе [3] рассматривается распределение, представляющее собой масштабную смесь распределения Рэлея с обобщенным гамма-распределением; приводятся выражения для его плотности, функции распределения, моментов, коэффициентов асимметрии и эксцесса, а также исследуются энтропийные свойства. На примере оценки реальных данных показана эффективность применения данной модели при анализе медицинских и надежности показателей. В статье [4] так называемые эрлангизированные масштабные смеси, представляющие собой смеси распределения Эрланга с распределением, имеющим неотрицательный носитель, применяются при построении нового метода вычисления вероятностей разорения в моделях риска Крамера–Лундберга и могут использоваться для анализа моделей, в которых распределения имеют тяжелые хвосты. В работах [5, 6] показана эффективность применения масштабной смеси распределения Фреше с гамма-распределением при исследовании реальных данных объема выпавших осадков. Ряд примеров можно продолжить.

В данной работе рассматриваются свойства функции, позволяющей описывать основные вероятностные характеристики масштабных смесей обобщенных гамма-распределений, к частным случаям которых относятся экспоненциальное распределение, гамма-распределение, χ - и χ^2 -распределения, распределения Эрланга, Накагами, Рэлея, Максвелла–Больцмана и Вейбулла, а также их масштабированные и обратные аналоги.

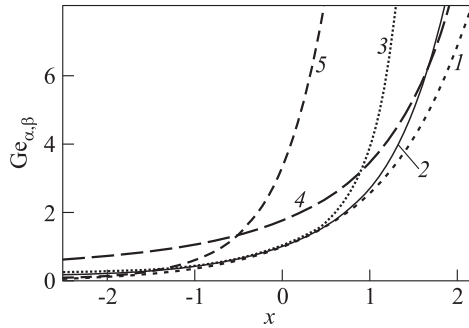
2 Гамма-экспоненциальная функция и ее аналитические свойства

В случае использования для моделирования распределений с неотрицательным неограниченным носителем широко используются частные случаи обобщенного гамма-распределения, имеющего плотность

$$f(x) = \frac{|v|x^{vq-1}e^{-(x/\theta)^v}}{\theta^{vq}\Gamma(q)}, \quad v \neq 0, \quad q > 0, \quad \theta > 0, \quad x > 0. \quad (1)$$

Распределение (1) впервые было предложено в 1920-х гг. итальянским экономистом Л. Аморозо [7], активно использовалось в 1940-х гг. советскими гидрологами С. Н. Крицким и М. Ф. Менкелем [8, 9], но зачастую ассоциируется с именем Е. Стэйси, рассмотревшего в 1962 г. частный случай распределения Аморозо [10].

Вероятностные характеристики масштабных смесей распределения (1) не представимы в элементарных функциях. В работе [11] для описания плотности произведения двух независимых случайных величин, имеющих распределения



Графики гамма-экспоненциальной функции при различных значениях параметров: 1 — $\alpha = 0,1, \beta = 1$; 2 — $\alpha = 0,5, \beta = 1$; 3 — $\alpha = 0,8, \beta = 1$; 4 — $\alpha = 0,5, \beta = 0,5$; 5 — $\alpha = 0,5, \beta = 3,5$

вида (1), было предложено использовать специальную гамма-экспоненциальную функцию

$$\text{Ge}_{\alpha, \beta}(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} \Gamma(\alpha k + \beta), \quad 0 \leq \alpha < 1, \quad \beta > 0, \quad x \in \mathbb{R}. \quad (2)$$

Функция (2) обобщает на случай $\beta \neq 1$ преобразование, введенное Леруа [12] для исследования производящих функций специального вида. Кроме того, функцию (2) можно рассматривать (при некоторых допущениях) как частный случай функции Сривастава–Томовски [13], обобщающей функцию Миттаг-Леффлера [14].

На рисунке приведены графики гамма-экспоненциальной функции при различных значениях параметров.

Аналитические свойства гамма-экспоненциальной функции во многом базируются на следующем утверждении [15].

Лемма 1. Пусть $a, b, p, s > 0$. Тогда

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

В работе [11] были обоснованы следующие аналитические свойства гамма-экспоненциальной функции (2). Везде далее, если не указано иное, следует считать, что утверждение выполнено для всех $\alpha \in [0, 1), \beta > 0, x \in \mathbb{R}, n = 0, 1, \dots$

1. $\text{Ge}_{\alpha, \beta}(x)$ — непрерывная функция.

2. $\text{Ge}_{1, n+1}(x) = (x^n/(1-x))_x^{(n)}$, $|x| < 1$. В частности, $\text{Ge}_{1, 1}(x) = 1/(1-x)$ и $\text{Ge}_{1, 2}(x) = 1/(1-x)^2$, $|x| < 1$. Под $\text{Ge}_{1, \beta}(x)$ в данном случае понимается ряд (2) при $\alpha = 1$.
3. $\text{Ge}_{\alpha, \beta}^{(n)}(x) = \text{Ge}_{\alpha, \alpha n + \beta}(x)$.
4. $\text{Ge}_{\alpha, \beta}(0) = \Gamma(\beta)$.
5. $\text{Ge}_{0, \beta}(x) = \Gamma(\beta)e^x$.
6. $\text{Ge}_{\alpha, 1}(x) = 1 + \alpha x \text{Ge}_{\alpha, \alpha}(x)$, $0 < \alpha < 1$.
7. $\text{Ge}_{\alpha, \beta+1}(x) = \alpha x \text{Ge}_{\alpha, \alpha+\beta}(x) + \beta \text{Ge}_{\alpha, \beta}(x)$.
8. $\sum_{k=0}^{\infty} (x^k/k!) \text{Ge}_{1/p, (k+p)/p}(-x) = 1$, $x > 0$, $p > 1$.
9. $\lim_{x \rightarrow -\infty} \text{Ge}_{\alpha, \beta}(x) = 0$.
10. $\int_0^{+\infty} \text{Ge}_{\alpha, \alpha+1}(-x) dx = 1$, $0 < \alpha < 1$.
11. $\text{Ge}_{\alpha, \beta}(x) > 0$.
12. $\text{Ge}_{\alpha, \beta}(x)$ — строго монотонная функция.

Базируясь на взаимосвязи основных вероятностных характеристик (плотность, функция распределения, моменты) масштабных смесей законов из гамма-класса [16], выражаемых в терминах функции (2), можно получить следующие утверждения.

13. $\int_0^x u^{p-1} \text{Ge}_{\alpha, \alpha p+1}(-u) du = \alpha x^p \text{Ge}_{\alpha, \alpha p}(-x)$, $p > 0$, $\alpha \in (0, 1)$, $x > 0$.
14. $\int_0^{\infty} u^{z+p-1} \text{Ge}_{\alpha, \alpha p+1}(-u) du = \Gamma(z+p)\Gamma(1-\alpha z)$, $p > 0$, $-p < z < \frac{1}{\alpha}$.
15. $\int_0^{\infty} u^z \text{Ge}_{\alpha, \alpha+\beta}(-u) du = \Gamma(z+1)\Gamma(\beta-\alpha z)$, $-1 < z < \frac{\beta}{\alpha}$.

Переходя к пределу при $x \rightarrow \infty$ в 13 и применяя 14, получаем

$$16. \lim_{x \rightarrow \infty} \alpha x^{\beta/\alpha} \text{Ge}_{\alpha, \beta}(-x) = \Gamma\left(\frac{\beta}{\alpha}\right), \quad \alpha \in (0, 1).$$

Из свойств 4 и 5, леммы 1 и теоремы Лебега о предельном переходе под знаком интеграла вытекает следующее утверждение:

$$17. \text{Ge}_{\alpha, \beta}(x) = \int_0^{\infty} y^{\beta-1} e^{-y+xy^{\alpha}} dy.$$

При помощи последнего свойства и теоремы о непрерывности несобственного интеграла по параметру нетрудно получить следующее утверждение:

18. $\text{Ge}_{\alpha, \beta}(x)$ непрерывна по параметрам α и β .

Следующее утверждение, доказанное в [12] для $\beta = 1$, допускает обобщение на случай произвольного $\beta > 0$ и обосновывается применением свойства 17 и теоремы о непрерывности несобственного интеграла по параметру:

$$19. \lim_{\alpha \rightarrow 1} \text{Ge}_{\alpha, \beta}(x) = \frac{\Gamma(\beta)}{1-x}, \quad x < 1.$$

Гамма-экспоненциальная функция допускает различные представления при помощи классических специальных функций. В работе [17] были обоснованы следующие утверждения:

$$20. \text{Ge}_{\alpha, \beta}(-x) = H_{11}^{11} \left[x \mid \begin{matrix} (\beta, \alpha) \\ (0, 1) \end{matrix} \right], \quad 0 < \alpha < 1, x > 0,$$

где $H_{11}^{11}[x \mid \cdot]$ — H -функция Фокса [18].

$$21. \text{Ge}_{m/n, c}(-x) = \sum_{k=0}^{n-1} \frac{(-x)^k}{k!} \Gamma \left(c + k \frac{m}{n} \right) {}_{m+1}F_n \left[1, \Delta \left(m, c + km/n \right); \Delta \left(n, 1 + k \right) \right] \times \\ \times \left(\frac{m^m}{n^n} (-x)^n \right), \quad c > 0, x > 0,$$

где m и n — взаимно простые натуральные числа ($m < n$); ${}_pF_q[\cdot](x)$ — обобщенная гипергеометрическая функция [18];

$$\Delta(k, a) = \frac{a}{k}, \frac{a+1}{k}, \dots, \frac{a+k-1}{k}.$$

Используя соотношение 5.2.11.14 из [19], получаем следующее утверждение:

$$22. \text{Ge}_{1/2, \beta}(x) = 2^{1-\beta} \Gamma(2\beta) e^{x^2/8} D_{-2\beta} \left(-\frac{x}{\sqrt{2}} \right),$$

где $D_{\nu}(z)$ — функция параболического цилиндра.

3 Алгоритм вычисления гамма-экспоненциальной функции

Алгоритм быстрого вычисления гамма-экспоненциальной функции представляет собой нетривиальную задачу. Следующие два утверждения играют в ней важную роль.

Обозначим модуль k -го члена ряда (2) через

$$g_k = \frac{|x|^k}{k!} \Gamma(\alpha k + \beta);$$

введем обозначения

$$\delta(k) \equiv \frac{g_{k+1}}{g_k} = \frac{|x|}{k+1} \frac{\Gamma(\alpha(k+1) + \beta)}{\Gamma(\alpha k + \beta)};$$

$$\bar{\delta}(k) = \frac{|x|}{k} \frac{\Gamma(\alpha(k+1) + \beta)}{\Gamma(\alpha k + \beta)}.$$

Теорема 1. Пусть $\alpha \in (0, 1)$, $\beta > 0$, $x \in \mathbb{R}$, $\varepsilon > 0$, $m \in \mathbb{N}$. Тогда если

$$\bar{\delta}(m) < 1, \quad \frac{g_m}{1 - \bar{\delta}(m)} < \varepsilon,$$

то

$$\sum_{k=m}^{\infty} \frac{|x|^k}{k!} \Gamma(\alpha k + \beta) < \varepsilon.$$

Доказательство. Покажем, что $\delta(y) < \bar{\delta}(m)$ при $y \geq m$. Из определения $\delta(y)$ и $\bar{\delta}(y)$ следует, что $\delta(y) < \bar{\delta}(y)$ при $y > 0$; следовательно, достаточно показать, что $\bar{\delta}(y)$ не возрастает на $y \in [1, +\infty)$. Имеем

$$\bar{\delta}'(y) = \frac{|x| \Gamma(\alpha(y+1) + \beta)}{y \Gamma(\alpha y + \beta)} \left(\alpha (\psi(\alpha(y+1) + \beta) - \psi(\alpha y + \beta)) - \frac{1}{y} \right),$$

где $\psi(z) = \Gamma'(z)/\Gamma(z)$ — дигамма-функция. Известно [19], что $\psi(z)$ строго возрастает на положительной полуоси, а также

$$\psi(z+1) = \psi(z) + \frac{1}{z}.$$

Отсюда

$$\alpha(\psi(\alpha(y+1) + \beta) - \psi(\alpha y + \beta)) < \alpha(\psi(\alpha y + \beta + 1) - \psi(\alpha y + \beta)) = \frac{\alpha}{\alpha y + \beta} < \frac{1}{y}$$

при $y > 0$ и, следовательно, $\bar{\delta}'(y) < 0$ при $y > 0$.

При $k \geq m$ для g_k имеем

$$g_k = \delta(k-1)\delta(k-2)\cdots\delta(m)g_m \leq (\bar{\delta}(m))^{k-m}g_m,$$

откуда

$$\begin{aligned} \sum_{k=m}^{\infty} \frac{|x|^k}{k!} \Gamma(\alpha k + \beta) &= \sum_{k=m}^{\infty} g_k \leq \sum_{k=m}^{\infty} (\bar{\delta}(m))^{k-m} g_m = \\ &= g_m \sum_{k=0}^{\infty} (\bar{\delta}(m))^k = \frac{g_m}{1 - \bar{\delta}(m)} < \varepsilon, \end{aligned}$$

что и требовалось доказать.

Теорема 2. Пусть $\alpha \in (0, 1)$, $\beta > 0$, $x \leq 0$, $\varepsilon > 0$, $m \in \mathbb{N}$. Тогда если $\bar{\delta}(m) < 1$, $g_m < \varepsilon$, то

$$\left| \sum_{k=m}^{\infty} \frac{x^k}{k!} \Gamma(\alpha k + \beta) \right| < \varepsilon.$$

Доказательство. В доказательстве теоремы 1 показано, что $\delta(k) < \bar{\delta}(m)$ при $k \geq m$. Следовательно, $\delta(k) < 1$ при $k \geq m$, откуда $g_{k+1} < g_k$ при $k \geq m$. Имеем

$$\left| \sum_{k=m}^{\infty} \frac{x^k}{k!} \Gamma(\alpha k + \beta) \right| = \left| \sum_{k=m}^{\infty} (-1)^k g_k \right| \leq g_m < \varepsilon.$$

Теорема доказана.

Основываясь на теоремах 1 и 2, можно предложить следующий алгоритм вычисления гамма-экспоненциальной функции.

При $\alpha = 0$ (см. свойство 5) $\text{Ge}_{0, \beta}(x) = \Gamma(\beta)e^x$. Задача вычисления ряда (2) при заданных $\alpha \in (0, 1)$, $\beta > 0$, $x \in \mathbb{R}$ и абсолютной точности $\varepsilon > 0$ сводится к нахождению такого m , что

$$\left| \sum_{k=m}^{\infty} \frac{x^k}{k!} \Gamma(\alpha k + \beta) \right| < \varepsilon.$$

Для обеспечения заданного уровня точности последовательно по $k = 0, 1, 2, \dots$ вычисляются члены ряда (2) и числа $\bar{\delta}(k-1)$, $k = 2, 3, \dots$. На каждом этапе проверяется выполнение условий теоремы 1, если $x > 0$, или теоремы 2, если $x < 0$. В случае выполнения требуемых условий процесс вычисления останавливается и выдается ответ.

При достаточно малых значениях $|x|$ данный алгоритм позволяет на порядки увеличить скорость вычисления гамма-экспоненциальной функции (2) по сравнению с методами численного интегрирования, основанными на свойстве 17. Этот эффект становится особенно заметен при увеличении требуемой точности вычислений.

4 Заключение

В работе представлены результаты, касающиеся свойств гамма-экспоненциальной функции. Эти результаты могут быть полезны при исследовании масштабных смесей вероятностных распределений, в первую очередь — обобщенных гамма-распределений. Приводимые аналитические свойства дополняют и расширяют известные утверждения для некоторых обобщений функции Mittag-Леффлера, а вычислительные алгоритмы позволяют использовать гамма-экспоненциальную функцию в приложениях.

Литература

1. Behme A., Bondesson L. A class of scale mixtures of Gamma(k)-distributions that are generalized gamma convolutions // *Bernoulli*, 2017. Vol. 23. No. 1. P. 773–787.
2. Korolev V. Yu., Gorshenin A. K., Zeifman A. I. On mixture representations for the generalized Linnik distribution and their applications in limit theorems // *J. Math. Sci.*, 2020. Vol. 246. No. 4. P. 503–518.
3. Rivera P. A., Barranco-Chamorro I., Gallardo D. I., Gómez H. W. Scale mixture of Rayleigh distribution // *Mathematics*, 2020. Vol. 8. Iss. 10. Art. ID: 1842. 22 p. doi: 10.3390/math8101842.
4. Peralta O., Rojas-Nandayapa L., Xie W., Yao H. Approximation of ruin probabilities via Erlangized scale mixtures // *Insur. Math. Econ.*, 2018. Vol. 78. P. 136–156.
5. Королев В. Ю., Горшенин А. К. О распределении вероятностей экстремальных осадков // *Докл. Акад. наук*, 2017. Т. 477. № 5. С. 604–609.
6. Gorshenin A. K., Korolev V. Yu. Scale mixtures of Frechet distribution as asymptotic approximations of extreme precipitation // *J. Math. Sci.*, 2018. Vol. 234. No. 6. P. 886–903.
7. Amoroso L. Ricerche intorno alla curva dei redditi // *Ann. Mat. Pur. Appl.* 1925. Vol. 21. P. 123–159.
8. Крицкий С. Н., Менкель М. Ф. О приемах исследования случайных колебаний речного стока // *Труды НИУ ГУГМС. Сер. IV*, 1946. Вып. 29. С. 3–32.
9. Крицкий С. Н., Менкель М. Ф. Выбор кривых распределения вероятностей для расчетов речного стока // *Известия АН СССР. Отд. техн. наук*, 1948. № 6. С. 15–21.
10. Stacy E. W. A generalization of the gamma distribution // *Ann. Math. Stat.*, 1962. Vol. 33. P. 1187–1192.
11. 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.
12. Le Roy É. Sur les séries divergentes et les fonctions définies par un développement de Taylor // *Ann. Faculté Sci. Toulouse 2 Sér.*, 1900. Vol. 2. No. 3. P. 317–384.
13. Srivastava H. M., Tomovski Ž. Fractional calculus with an integral operator containing a generalized Mittag-Leffler function in the kernel // *App. Math. Comput.*, 2009. Vol. 211. P. 198–210.
14. Gorenflo R., Kilbas A. A., Mainardi F., Rogosin S. V. Mittag-Leffler functions, related topics and applications. — Berlin, Heidelberg: Springer-Verlag, 2014. 443 p.

15. *Кудрявцев А. А., Палионная С. И., Шоргин В. С.* Априорное обратное гамма-распределение в байесовских моделях массового обслуживания // Системы и средства информатики, 2018. Т. 28. Вып. 4. С. 52–58.
16. *Арутюнов Е. Н., Кудрявцев А. А., Титова А. И.* Гамма-вейбулловский случай априорных распределений в байесовских моделях массового обслуживания // Информатика и её применения, 2018. Т. 12. Вып. 4. С. 92–95.
17. *Арутюнов Е. Н., Кудрявцев А. А., Недоливко Ю. Н.* Вероятностные характеристики индекса баланса факторов, имеющих обобщенные гамма-распределения // Информатика и её применения, 2021. Т. 15. Вып. 1. С. 11–17.
18. *Прудников А. П., Брычков Ю. А., Маричев О. И.* Интегралы и ряды: в 3 т. Т. 3: Специальные функции. Дополнительные главы. — 2-е изд. — М.: Физматлит, 2003. 688 с.
19. *Прудников А. П., Брычков Ю. А., Маричев О. И.* Интегралы и ряды: в 3 т. Т. 1: Элементарные функции. — 2-е изд. — М.: Физматлит, 2002. 632 с.

Поступила в редакцию 15.03.21

ANALYTICAL PROPERTIES AND ASPECTS OF COMPUTATION OF THE GAMMA-EXPONENTIAL FUNCTION

M. O. Vorontsov¹, A. A. Kudryavtsev¹, and S. Ya. Shorgin²

¹Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskie 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: Mixtures of probability distributions play an important role in modern analysis and modeling of complex processes. Traditionally, much research attention is paid to the distributions from the gamma class. The main probabilistic characteristics of scale mixtures of generalized gamma distributions cannot be expressed in elementary functions that complicates the process of analytical research and often leads to unreasonably large computational difficulties. The paper analyzes aspects of computation and analytical properties of the gamma-exponential function which has proven itself as a convenient tool for studying scale mixtures of generalized gamma distributions. The presented results expand the usage of functions like the Mittag-Leffler function in the analysis of probability distributions.

Keywords: gamma-exponential function; generalized gamma distribution; scale mixture; computational algorithms

DOI: 10.14357/08696527210210

Acknowledgments

The work was partly supported by the Russian Foundation for Basic Research (project 20-07-00655). The research was conducted in accordance with the program of the Moscow Center for Fundamental and Applied Mathematics.

References

1. Behme, A., and L. Bondesson. 2017. A class of scale mixtures of Gamma(k)-distributions that are generalized gamma convolutions. *Bernoulli* 23(1):773–787.
2. Korolev, V. Yu., A. K. Gorshenin, and A. I. Zeifman. 2020. On mixture representations for the generalized linnik distribution and their applications in limit theorems. *J. Math. Sci.* 246(4):503–518.
3. Rivera, P. A., I. Barranco-Chamorro, D. I. Gallardo, and H. W. Gómez. 2020. Scale mixture of Rayleigh distribution. *Mathematics* 8(10):1842. 22 p. doi: 10.3390/math8101842.
4. Peralta, O., L. Rojas-Nandayapa, W. Xie, and H. Yao. 2018. Approximation of ruin probabilities via Erlangized scale mixtures. *Insur. Math. Econ.* 78:136–156.
5. Korolev, V. Yu., and A. K. Gorshenin. 2017. The probability distribution of extremal precipitations. *Dokl. Earth Sci.* 477(2):1461–1466.
6. Gorshenin, A. K., and V. Yu. Korolev. 2018. Scale mixtures of Frechet distribution as asymptotic approximations of extreme precipitation. *J. Math. Sci.* 234(6):886–903.
7. Amoroso, L. 1925. Ricerche intorno alla curva dei redditi. *Ann. Mat. Pur. Appl.* 21:123–159.
8. Kritsky, S. N., and M. F. Menkel. 1946. O priemakh issledovaniya sluchaynykh kolebaniy rechnogo stoka [Methods of investigation of random fluctuations of river flow]. *Trudy NIU GUGMS Ser. IV* [Proceedings of GUGMS Research Institutions, Ser. IV] 29:3–32.
9. Kritsky, S. N., and M. F. Menkel. 1948. Vybor krivyykh raspredeleniya veroyatnostey dlya raschetov rechnogo stoka [Selection of probability distribution curves for river flow calculations]. *Izvestiya AN SSSR. Otd. tekhn. nauk* [Herald of the Russian Academy of Sciences. Technical Sciences] 6:15–21.
10. Stacy, E. W. 1962. A generalization of the gamma distribution. *Ann. Math. Stat.* 33:1187–1192.
11. 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.
12. Le Roy, É. 1900. Sur les séries divergentes et les fonctions définies par un développement de Taylor. *Ann. Faculté Sci. Toulouse 2 Sér.* 2(3):317–384.
13. Srivastava, H. M., and Ž. Tomovski. 2009. Fractional calculus with an integral operator containing a generalized Mittag-Leffler function in the kernel. *Appl. Math. Comput.* 211:198–210.
14. Gorenlo, R., A. A. Kilbas, F. Mainardi, and S. V. Rogosin. 2014. *Mittag-Leffler functions, related topics and applications*. Berlin–Heidelberg: Springer-Verlag. 443 p.

15. Kudryavtsev, A. A., S. I. Palionnaia, and V. S. Shorgin. 2018. Apriornoe obratnoe gamma-raspredelenie v bayesovskikh modelyakh massovogo obsluzhivaniya [*A priori inverse gamma distribution in Bayesian queuing models*]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(4):52–58.
16. Arutyunov, E. N., A. A. Kudryavtsev, and A. I. Titova. 2018. Gamma-veybullovskiy sluchay apriornykh raspredeleniy v bayesovskikh modelyakh massovogo obsluzhivaniya [*Gamma-Weibull a priori distributions in Bayesian queuing models*]. *Informatika i ee Primeneniya — Inform. Appl.* 12(4):92–95.
17. Arutyunov, E. N., A. A. Kudryavtsev, and Iu. N. Nedolivko. 2021. Veroyatnostnye kharakteristiki indeksa balansa faktorov, imeyushchikh obobshchennye gamma-raspredeleniya [Probabilistic characteristics of balance index of factors with generalized gamma distribution]. *Informatika i ee Primeneniya — Inform. Appl.* 15(1):11–17.
18. Prudnikov, A. P., Yu. A. Brychkov, and O. I. Marichev. 2003. *Integraly i ryady. V 3 t. T. 3. Spetsial'nye funktsii. Dopolnitel'nye glavy* [Integrals and series. In 3 vols. Vol. 3. Special functions. Additional chapters]. 2nd ed. Moscow: Fizmatlit. 688 p.
19. Prudnikov, A. P., Yu. A. Brychkov, and O. I. Marichev. 2002. *Integraly i ryady. V 3 t. T. 1. Elementarnye funktsii* [Integrals and series. In 3 vols. Vol. 3. Elementary functions]. 2nd ed. Moscow: Fizmatlit. 632 p.

Received March 15, 2021

Contributors

Vorontsov Mikhail O. (b. 1996) — student, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation; m.vtsov@mail.ru

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 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation; nubigena@mail.ru

Shorgin Sergey Ya. (b. 1952) — Doctor of Science in physics and mathematics, professor, principal scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; sshorgin@ipiran.ru

ВАРИАНТ КОНЦЕПТУАЛЬНОЙ СХЕМЫ БАЗЫ ГЕОДАНЫХ С ПОДДЕРЖКОЙ ОБРАТИМОЙ ГЕНЕРАЛИЗАЦИОННОЙ СВЯЗАННОСТИ МОДЕЛЕЙ ГЕООБЪЕКТОВ

*Д. А. Никишин*¹

Аннотация: Представлена структура концептуальной схемы и основные проектные решения для построения варианта базы геоданных (БГД), обеспечивающей поддержку обратимой генерализационной связанности геомоделей, представляющих геообъекты с разной степенью детализации. Это позволит расширить функциональность перспективных БГД за счет применения мультимасштабных методов анализа геоданных и решения по ним соответствующих задач. В частности, такая мультимасштабная структура БГД позволит формализовать процессы генерализации и контроля целостности данных в БГД и тем способствовать их автоматизации.

Ключевые слова: базы геоданных; концептуальная схема базы геоданных; моделирование геообъектов; обратимая генерализация геомоделей

DOI: 10.14357/08696527210211

1 Введение

Ранее, в публикации [1], были рассмотрены предпосылки для создания единой концептуальной схемы² классификации топографической информации, нацеленной в том числе на унификацию и интеграцию частных концептуальных схем для отдельных уровней детализации (масштабов) и прикладных задач. В этом контексте в данной работе детально рассмотрены основные проектные решения и структуры данных, предназначенные для построения возможного варианта модели БГД, поддерживающей обратимую генерализационную связанность отдельных моделей, представляющих геообъекты с разной степенью детализации. Это позволит обеспечить потенциал для совершенствования методологии анализа и отображения информации о местности, которую можно рассматривать как один из важнейших компонентов информационных, управляющих и телекоммуникационных систем (ИУТС).

Исследования проводятся в контексте общего направления исследований информационных трансформаций [2] и полиадического компьютеринга [3] в рамках

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

²Концептуальная схема здесь понимается как система взаимосвязанных понятий, необходимая и достаточная для описания требуемого аспекта моделируемого объекта, включающая модели геообъектов, сопряженные с ними ограничения и методы их обработки.

задачи исследования свойства обратимости процессов генерализации пространственно-временных данных в ИУТС.

2 Множественность представления геообъектов

В БГД предполагается наличие нескольких уровней представления данных о геообъектах, которые отличаются между собой степенью обобщения (генерализации), или же, от противного, — степенью подробности, детальности данных [4]. Для обозначения такого уровня используется конвенциональный термин — *уровень детализации* (англ. Level of Details, LoD) [5, 6].

Геообъект в зависимости от своих размеров на местности и текущего уровня детализации модели местности (ее «масштабного уровня») может описываться различно: как составной объект [7], как однородная область (полигон) либо как обобщенный, «внемасштабный» объект (линейный или точечный). При этом различия этих представлений (моделей) геообъекта носят комплексный характер: речь идет об обобщении не только пространственного описания, но и семантических характеристик геообъекта. Кроме того, геообъект может иметь разные представления даже в контексте одного LoD, также отличающиеся по степени детальности пространственного и/или семантического описания [1].

Таким образом, каждый геообъект в общем случае могут представлять несколько вариантов его модельных описаний (моделей геообъекта), соответствующих разным LoD, для краткости эти описания далее будем именовать *геомоделями*. На рис. 1 показан пример возможного набора вариантов геомodelей для некоторого геообъекта (знаком «+» обозначено наличие соответствующих геомodelей). Разным колонкам соответствуют геомodelи для разных LoD, а в пределах одной колонки (одного LoD) присутствуют альтернативные модели, различающиеся по типу метрики и/или семантическому описанию.

В целях обеспечения согласованного использования всех этих моделей в процессах анализа и обработки данных должна соблюдаться преемственность представлений геообъекта как в смежных LoD, так и между различными моделями в пределах одного уровня обобщения [1]. Такой подход призван обеспечить совершенствование функциональности методов анализа геоданных [8,9] и решения других задач [10–12].

Тип геомodelи	Уровень детализации			
	Scale 1	Scale 2	...	Scale N
Aggregate	+		...	
SQR	+	+	...	
LIN		+	...	
DOT, VEC			...	+

Рис. 1 Пример набора вариантов геомodelей

Для соблюдения такой преемственности и расширения функциональности БГД важным принципом представляется обеспечение *генерализационной связанности* — наличия явно определенных связей между представлениями геообъектов, соответствующих смежным LoD [4]. При этом необходимо применение принципа *обратимости* генерализации [13], сущность которой заключается в обеспечении двунаправленных взаимосвязей между представлениями геообъекта на смежных LoD. Иерархическая структура «пирамиды» генерализации и типология связей в ней рассматривались в работе [14].

Генерализационная связанность, обеспечивая прямое обращение к представлениям геообъекта на разных LoD, позволит реализовать методы *многоуровневого* поиска, анализа и обработки данных в БГД (включая задачи верификации и визуализации геоданных), что обеспечит потенциал для совершенствования функций геоинформационных систем (ГИС). Примером может служить обеспечение углубленного доступа к подробным семантическим характеристикам и компонентам геообъектов, присущим более детальным уровням моделирования при создании тематических представлений ситуации для специфических задач.

Так, непроезжий участок улицы в мелком масштабе описывается слишком обобщенно, без определения характера препятствия для проезда. На детальных уровнях может иметь место дифференциация: значительный уклон (крутой склон), ступенчатый продольный профиль (лестница) либо препятствия типа валунов, бетонных блоков и т. п. Соответственно, отсутствие такой дифференциации делает невозможным определение проходимости такого участка для отдельных видов транспортных средств или применимость того или иного способа его преодоления в мелком масштабе без обращения к крупномасштабным данным.

Применение генерализационной связанности также может служить решению проблемы автоматической каскадной генерализации — в этом случае достаточно будет осуществить контролируемую генерализацию только один раз — для вновь появившихся объектов, а изменения существующих объектов, вносимые на нижнем уровне, должны автоматически распространяться на вышестоящие уровни генерализации.

При отсутствии в БГД явных генерализационных связей задачи анализа данных требуют сложных процедур поиска и извлечения геомodelей требуемой полноты и детальности в наборах данных соответствующего LoD.

3 Модель мультимасштабной базы геоданных

В данном разделе на рис. 2 и 3 представлена¹ структура концептуальной схемы БГД, являющаяся попыткой практической реализации БГД с поддержкой обратимой генерализационной связанности моделей, представляющих геообъект

¹ В целях более четкого выделения структур данных и сокращения объема на рис. 2 и 3 некоторые элементы структур опущены или оформлены в упрощенном виде, с применением псевдокода.

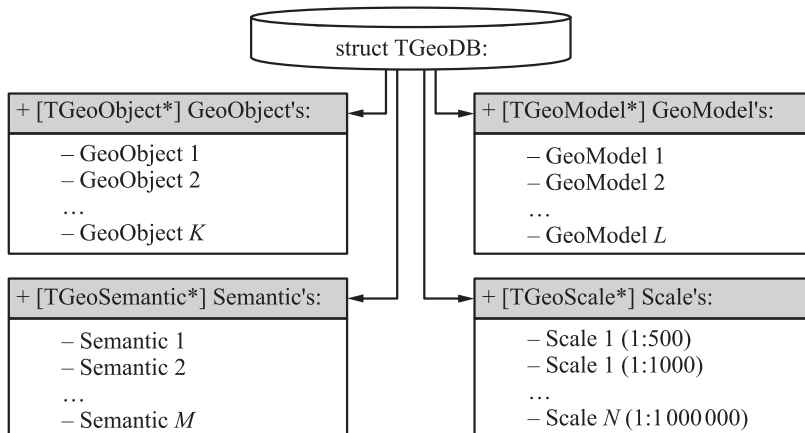


Рис. 2 Общая структура МБГД

с разной степенью детализации. Такую БГД далее условно будем именовать мультимасштабной БГД (МБГД).

Данная концептуальная схема является прототипом для рассмотренной в [1] более общей БГД, включающей в себя также поддержку вариантного и темпорального аспектов геоданных.

На рис. 2 представлена общая структура МБГД, обозначенная как TGeoDB. Она включает в себя четыре основные таблицы, по своей структуре соответствующие четырем основным концептам МБГД: «геообъект» (TGeoObject), «геомодель» (TGeoModel), «семантическое свойство» (TGeoSemantic) и «масштабный уровень» (TGeoScale); структура этих концептов рассмотрена ниже. Данные таблицы содержат собственно данные БГД и представляют собой основные точки входа, предназначенные для обращения к соответствующим информационным объектам геоданных.

На рис. 3 представлена структура основных концептов МБГД и связей между ними. Рассмотрим их подробнее, при этом концепты будем рассматривать как классы (структуры) данных соответствующего информационного объекта.

1. Класс TGeoObject представляет собой структуру данных, описывающую конкретный объект местности. Она включает поле наименования геообъекта и указатели на все имеющиеся для этого объекта геомодели (модели геообъекта) в разных масштабах (LoD) и/или для разных типов метрики.
2. Класс TGeoScale — это структура, которая представляет данные конкретного уровня детализации («масштабного уровня») в МБГД. Она включает в себя поле Scale, описывающее номинальное значение масштаба (конкретная форма его представления — строкой либо числом — принципиального значения не

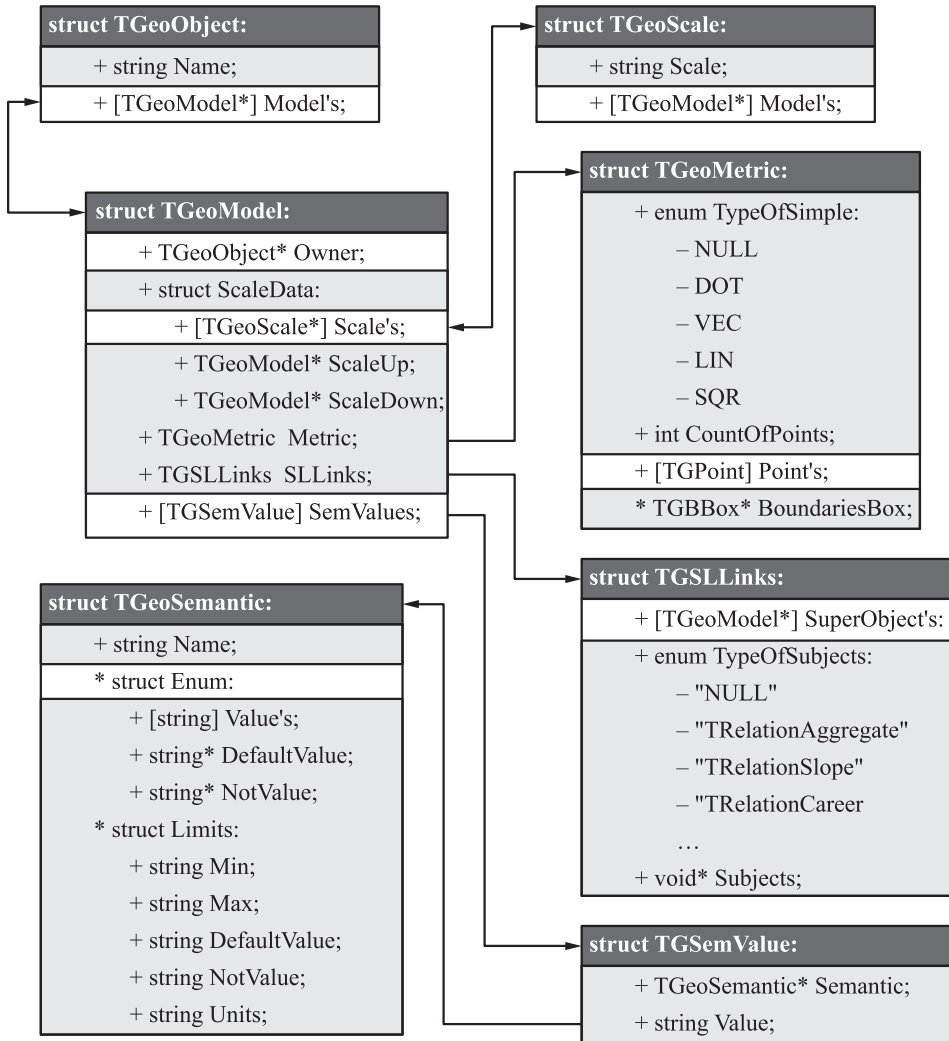


Рис. 3 Основные структуры данных для МБГД

имеет), а также совокупность указателей на геообъекты, релевантные данному LoD.

3. Класс TGeoModel — основной тип данных в МБГД, предназначенный для описания геообъекта; он включает следующие компоненты:

- (а) поле Owner представляет собой указатель на геообъект, который описывается данной геообъектом;

- (б) структуру `ScaleData`, которая, по сути, образует мультимасштабную структуру МБГД; она включает: совокупность указателей `Scale's`, определяющую масштабный диапазон, в котором применима данная геодезь; свойства `ScaleUp` и `ScaleDown` указывают на другие геодезь данного геообъекта, более обобщенную и более детальную соответственно, при этом масштабные диапазоны геодезь не должны перекрываться;
- (в) структуру `Metric`, которая предназначена для описания пространственной локализации геодезь элементарными типами метрики. Однако ее наличие предусмотрено и для составной геодезь, где наряду с компонентами может быть также предусмотрено пространственное описание в целом как результат генерализации (например, полигон квартала в дополнение к отдельным строениям). Структура `Metric` включает:
- перечисление `TypeOfSimple`, которое определяет конкретный тип метрики данной геодезь: значение `NULL` означает отсутствие данных о пространственной локализации; остальные значения соответствуют конвенциональным типам метрики (пространственного описания) геообъектов: `DOT` (точечный), `VEC` (точка + направление), `LIN` (полилиния)¹, `SQR` (полигон — область с определенными внутренней и внешней областями);
 - значение `CountOfPoints` для набора точек-узлов метрики `Point's`, которое частично определяется типом метрики (1 для `DOT`; 2 для `VEC`; ≥ 2 для всех остальных);
 - структуру `BoundariesBox`, которая может быть предусмотрена в целях поддержки некоторых процедур визуализации и анализа геодезь;
- (г) структуру `SLLinks`, которая описывает пространственно-логические связи (ПЛС) составной геодезь с ее компонентами. При этом имеется в виду, что составная геодезь не должна дублировать свойства ее компонентов (аналогично подходу, принятому в модели `OpenStreetMap`). Эта структура включает следующие элементы:
- `SuperObject's`, который представляет собой набор указателей на геодезь, использующие данную в качестве компонента. В этом качестве геодезь одновременно может принадлежать нескольким составным объектам;
 - перечисление `TypeOfSubjects`, которое определяет тип составной геодезь следующими значениями: `NULL` означает отсутствие компонентов у данной геодезь (т.е. она является элементарной); `TRelationAggregate` представляет агрегацию — совокупность компонентов без выделенных ролей и прочих ограничений; `TRelationSlope`,

¹При этом предполагаются известные соглашения для полилинейной метрики: последовательность ее вершин определяет направление, а совпадение начальной и конечной точек — ее замкнутость.

TRelationCareer и др. обозначают соответствующие разновидности комплексных геомodelей¹;

- поле Subjects, которое указывает на структуру соответствующего TypeOfSubjects типа, описывающую составную геомodelь посредством ПЛС. Некоторые разновидности данных структур были представлены в [7];
 - (д) структуру SemValues, которая содержит набор значений семантических свойств геомodelи. Каждый ее элемент типа TGSemValue включает указатель на объект типа TGeoSemantic, описывающий определенное семантическое свойство, а также строковое поле Value, используемое для задания значения этого свойства в конкретной геомodelи.
4. Класс TGeoSemantic описывает определенное семантическое свойство, имеющееся в БГД. Он включает наименование (Name), а также предусматривает два возможных варианта описания множества значений:
- (а) как перечислимого множества возможных значений свойства;
 - (б) как области числовых значений, ограниченной значениями Min и Max; для этого варианта предусмотрено описание единиц Units (для безразмерных должно иметь значение NULL).

В обоих вариантах предусмотрены также элементы DefaultValue (значение по умолчанию, в том числе NULL при отсутствии такового) и NotValue — «запрещенное значение», означающее отсутствие какого-либо определенного значения для данного свойства (что означает неприменимость данного свойства к конкретной геомodelи).

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

4 Заключение

В работе представлена структура концептуальной схемы и основные проектные решения для построения образца БГД, обеспечивающей поддержку согласованного мультимасштабного представления геообъектов за счет обратимой генерализационной связанности геомodelей, представляющих геообъекты с разной степенью детализации. Это потенциально позволит расширить функциональность БГД за счет применения мультимасштабных методов поиска, анализа и обработки геоданных и решения по ним соответствующих задач. В частности, обеспечение генерализационной связанности в БГД позволит формализовать процессы генерализации и контроля целостности данных в БГД и тем способствовать их автоматизации. Обеспечение автоматизации этих процессов жизненно

¹В которых определены роли для компонентов, количества компонент в каждой из них, а также некоторые другие правила.

необходимо для современных ГИС, к которым предъявляются требования оперативной актуализации данных, их анализа, обработки и визуализации, в том числе в реальном масштабе времени. В перспективе предполагается дополнить предлагаемую модель БГД возможностью вариантного представления геообъектов, а также добавить поддержку темпорального аспекта.

Литература

1. Дулин С. К., Никишин Д. А. Подходы к интеграции прикладных концептуальных схем в составе унифицированной геоонтологии // Системы и средства информатики, 2020. Т. 30. № 2. С. 68–77.
2. Зацман И. М. Методология обратимой генерализации в контексте классификации информационных трансформаций // Системы и средства информатики, 2018. Т. 28. № 2. С. 128–144.
3. Rosenbloom P. S. On computing: The fourth great scientific domain. — Cambridge, MA, USA: MIT Press, 2013. 307 p.
4. Никишин Д. А. Структура и особенности генерализации в контексте функционирования темпоральной базы геоданных // Системы и средства информатики, 2020. Т. 30. № 1. С. 147–159.
5. Verdie Y., Lafarge F., Alliez P. LOD generation for urban scenes // ACM T. Graphics, 2015. Vol. 34. Iss. 3. Art. No. 30. 14 p. doi: 10.1145/2732527.
6. Trajcevski G., Balasubramani B. S., Cruz I. F., Tamassia R., Teng X. Semantically augmented range queries over heterogeneous geospatial data // 28th Conference (International) on Advances in Geographic Information Systems. — New York, NY, USA: ACM, 2020. P. 68–77. doi: 10.1145/3397536.3422271.
7. Никишин Д. А. Подход к совершенствованию концептуальных схем баз геоданных посредством моделей для пространственно-логического связывания геообъектов // Системы и средства информатики, 2021. Т. 31. № 1. С. 157–167.
8. Pandey V., Kipf A., Vorona D., Mühlbauer T., Neumann T., Kemper A. High-performance geospatial analytics in HyPerSpace // Conference (International) on Management of Data Proceedings. — New York, NY, USA: ACM, 2016. P. 2145–2148. doi: 10.1145/2882903.2899412.
9. Wang S., Zhong Y., Lu H., Wang E., Yun W., Cai W. Geospatial big data analytics engine for spark // 6th ACM SIGSPATIAL Workshop on Analytics for Big Geospatial Data Proceedings. — New York, NY, USA: ACM, 2017. P. 42–45. doi: 10.1145/3150919.3150923.
10. Polys N. F., Singh A., Sforza P. A novel level-of-detail technique for virtual city environments // 21st Conference (International) on Web3D Technology Proceedings. — New York, NY, USA: ACM, 2016. P. 183–184. doi: 10.1145/2945292.2945322.
11. Wang H., Chen X., Polys N., Sforza P. A Web3D forest geo-visualization and user interface evaluation // 22nd Conference (International) on 3D Web Technology Proceedings. — New York, NY, USA: ACM, 2017. Art. No. 9. 9 p. doi: 10.1145/3055624.3075956.
12. Vollmer J. O., Trapp M., Schumann H., Döllner J. Hierarchical spatial aggregation for level-of-detail visualization of 3D thematic data // ACM T. Spatial Algorithms Systems, 2018. Vol. 4. Iss. 3. Art. No. 9. 23 p. doi: 10.1145/3234506.

13. *Никишин Д. А.* Сопоставление особенностей представления геоданных в картографии и геоинформатике // Системы и средства информатики, 2018. Т. 28. № 2. С. 60–74.
14. *Никишин Д. А.* Виды неоднородностей в структуре генерализации геоданных // Системы и средства информатики, 2019. Т. 29. № 1. С. 74–85.

Поступила в редакцию 03.03.21

A VARIANT OF THE CONCEPTUAL SCHEMA OF THE GEODATA DATABASE WITH SUPPORT FOR REVERSIBLE GENERALIZATION CONNECTIVITY OF GEO OBJECT MODELS

D. A. Nikishin

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

Abstract: The structure of the conceptual scheme and the main design solutions for constructing a variant of the geodata database (GDB), which provides support for reversible generalization connectivity of geomodels representing a geo object with different details, are presented. This will expand the functionality of advanced GDB by using multiscale methods of geodata analysis and solving the corresponding tasks based on them. In particular, such a multiscale structure of the GDB will formalize the processes of generalization and automated control of data integrity in the GDB, and thus contribute to their automation.

Keywords: geodata databases; conceptual schema of the geodata database; modeling of geo objects; reversible generalization of geomodels

DOI: 10.14357/08696527210211

References

1. Dulin, S. K., and D. A. Nikishin. 2020. Podkhody k integratsii prikladnykh kontseptual'nykh skhem v sostave unifitsirovannoy geontologii [Approaches to application integration conceptual schemas in the composition of the unified geontologies]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(2):68–77.
2. Zatsman, I. M. 2018. Metodologiya obratimoy generalizatsii v kontekste klassifikatsii informatsionnykh transformatsiy [Methodology of reversible generalization in the context of classification of information transformations]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(2):128–144.
3. Rosenbloom, P. S. 2013. *On computing: The fourth great scientific domain*. Cambridge, MA: MIT Press. 307 p.
4. Nikishin, D. A. 2020. Struktura i osobennosti generalizatsii v kontekste funktsionirovaniya temporal'noy bazy geodannykh [Structure and features of generalization in the

- context of functioning of a temporal geodatabase]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(1):147–159.
5. Verdie, Y., F. Lafarge, and P. Alliez. 2015. LOD generation for urban scenes. *ACM T. Graphics* 34(3):30. 14 p. doi: 10.1145/2732527.
 6. Trajcevski, G., B. S. Balasubramani, I. F. Cruz, R. Tamassia, and X. Teng. 2020. Semantically augmented range queries over heterogeneous geospatial data. *28th Conference (International) on Advances in Geographic Information Systems Proceedings*. ACM. 68–77. doi: 10.1145/3397536.3422271.
 7. Nikishin, D. A. 2021. Podkhod k sovershenstvovaniyu kontseptual'nykh skhem baz geodannykh posredstvom modeley dlya prostranstvenno-logicheskogo svyazyvaniya geoob'ektov [An approach to improving the conceptual schemes of geodata by means of models for spatial and logical linking of geobjects]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 31(1):157–167.
 8. Pandey, V., A. Kipf, D. Vorona, T. Mühlbauer, T. Neumann, and A. Kemper. 2016. High-performance geospatial analytics in HyPerSpace. *2016 Conference (International) on Management of Data Proceedings*. New York, NY: ACM. 2145–2148. doi: 10.1145/2882903.2899412.
 9. Wang, S., Y. Zhong, H. Lu, E. Wang, W. Yun, and W. Cai. 2017. Geospatial big data analytics engine for Spark. *6th ACM SIGSPATIAL Workshop on Analytics for Big Geospatial Data Proceedings*. New York, NY: ACM. 42–45. doi: 10.1145/3150919.3150923.
 10. Polys, N. F., A. Singh, and P. Sforza. 2016. A novel level-of-detail technique for virtual city environments. *21st Conference (International) on 3D Web Technology Proceedings*. New York, NY: ACM. 183–184. doi: 10.1145/2945292.2945322.
 11. Wang, H., X. Chen, N. Polys, and P. Sforza. 2017. A Web3D forest geo-visualization and user interface evaluation. *22nd Conference (International) on 3D Web Technology Proceedings*. New York, NY: ACM. Art. No. 9. 9 p. doi: 10.1145/3055624.3075956.
 12. Vollmer, J. O., M. Trapp, H. Schumann, and J. Döllner. 2018. Hierarchical spatial aggregation for level-of-detail visualization of 3D thematic data. *ACM T. Spatial Algorithms Systems* 4(3):9. 23 p. doi: 10.1145/3234506.
 13. Nikishin, D. A. 2018. Sopostavlenie osobennostey predstavleniya geodannykh v kartografii i geoinformatike [Comparison of characteristics of the representation of geodata in cartography and geoinformatics]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(2):60–74.
 14. Nikishin, D. A. 2019. Vidy neodnorodnostey v strukture generalizatsii geodannykh [Types of inhomogeneities in the structure of geodata generalization]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(1):74–85.

Received March 3, 2021

Contributor

Nikishin Dmitry A. (b. 1976)— Candidate of Science (PhD) in technology, 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; dmnikishin@mail.ru

НАСЛЕДУЕМЫЕ ЛЕКСИКОГРАФИЧЕСКИЕ РЕСУРСЫ БАЗЫ ДАННЫХ ФРАЗЕОЛОГИЧЕСКОГО СЛОВАРЯ

В. В. Вакуленко¹, И. М. Зацман²

Аннотация: Рассматриваются принципы импорта в базу данных (БД) фразеологического словаря ранее созданных словарных статей, т. е. использование наследуемых лексикографических ресурсов. Импорт словарных статей в БД основан на двух основных принципах: структуризации текста статей и дополнении словарных статей новыми полями и их значениями, отсутствующими в статьях. Каждая статья включает идиому немецкого языка, один или несколько ее переводов на русский язык, их контексты и ряд других компонентов. Цель статьи состоит в описании принципов использования и пополнения наследуемых лексикографических ресурсов в процессе формирования БД фразеологического словаря немецких идиом, а также в сопоставлении функционала БД с возможностями аналогичных лексикографических информационных систем создания словарей в аспекте двух принципов импорта словарных статей.

Ключевые слова: принципы импорта; наследуемые лексикографические ресурсы; сопоставительный анализ; база данных фразеологического словаря

DOI: 10.14357/08696527210212

1 Введение

К основным задачам, которые решаются при помощи лексикографических систем и БД статей моно- и двуязычных словарей, относятся [1–4]:

- (1) импорт ранее созданных словарных статей, т. е. использование наследуемых лексикографических ресурсов;
- (2) создание новых словарных статей для формирования словарей;
- (3) редактирование словарных статей;
- (4) ведение классификационных систем и рубрицирование словарных статей, например отнесение статьи к одному или нескольким семантическим классам (время, пространство, движение, вещество и т. д.) и подклассам, которые могут образовывать многоуровневую иерархию [5];

¹Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, vvak@ipm.me

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

- (5) добавление гиперссылок между двумя и более словарными статьями или их компонентами;
- (6) поиск по тексту всей статьи или отдельным ее компонентам (например, только по примерам перевода в статьях двуязычных словарей), визуализация найденных словарных статей;
- (7) формирование словарей для издания и переиздания в бумажной и электронной форме в режиме экспорта статей из БД.

Цель статьи состоит в описании принципов использования наследуемых лексикографических ресурсов в процессе формирования БД фразеологического словаря немецких идиом и их переводов на русский язык, т. е. она относится к решению первой из семи перечисленных задач, а также в сопоставлении функционала БД с возможностями аналогичных лексикографических информационных систем создания словарей.

2 Принципы импорта

Используемые наследуемые ресурсы сформированы коллективом лексикографов под руководством доктора филологических наук Д. О. Добровольского. Они представляют собой массив словарных статей в формате редактора Word, где каждая статья есть отдельный текстовый документ, состоящий из нескольких компонентов. При этом на стадии подготовки словаря могут быть сформированы несколько вариантов одной статьи, что необходимо учитывать при их импорте в БД.

Пример 1 включает следующие шесть компонентов словарной статьи: (1) вокабула **SINTFLUT** (*нем.* потоп), согласно которой определяется порядок следования статей в бумажном словаре; (2) идиома на немецком языке; (3) ее перевод на русский язык; (4) стилистическая помета (*нейтр.*); (5) примеры использования идиомы с переводом (начало каждого примера обозначено символом ¶); (6) инициалы автора статьи Владимира Карпова.

Пять из шести перечисленных компонентов (кроме стилистических помет) являются обязательными и присутствуют во всех словарных статьях. В них факультативно могут также присутствовать, например, варианты перевода идиомы и комментарии. Отдельные компоненты могут состоять из зон. Каждый перевод в примере 1 включает три зоны: идиома в оригинальном контексте, его источник и перевод контекста на русский язык. Зона источника включает два поля: его название и дату.

Импорт словарных статей в БД основан на двух основных принципах: (1) структуризации текста статей; (2) дополнении словарных статей новыми полями и их значениями, отсутствующими в наследуемых ресурсах.

Пример 1 Компоненты словарной статьи

SINTFLUT

nach mir [uns] die Sintflut

после меня [нас] хоть потоп

neutr.

☞ Es ist ein wichtiges Anliegen von Eltern, ihren Nachkommen etwas zu hinterlassen und nicht zu sagen: *Nach mir die Sintflut.* (Mannheimer Morgen, 10.11.2007)

Важная задача родителей — оставить что-то своим детям в наследство, а не действовать по принципу «*после меня хоть потоп*».

☞ Insbesondere die ungeklärte Entsorgungsfrage zeigt, diese Leute sehen nicht ein, dass wir diesen Planeten nur von unseren Kindern und Enkeln geliehen haben! Stattdessen verfahren sie weiter, gemäß dem Motto: *Nach uns die Sintflut!* (Mannheimer Morgen, 12.09.2009)

В особенности нерешенность вопроса утилизации показывает: люди просто не осознают, что эту планету мы только взяли в аренду у наших детей и внуков! Вместо этого они продолжают жить по принципу «*после нас хоть потоп*».

БК

2.1 Первый принцип

Структуризация текста словарных статей, которая выполняется до их загрузки в БД, основана на трех правилах создания статей лексикографами:

- (1) соблюдение порядка размещения компонентов, зон и полей в тексте статьи, а также *нумерованных значений* идиомы немецкого языка (пример 2);
- (2) включение в текст статей дат в определенном формате, символов начала примеров, зон вариантов и комментариев, а также спецсимволов (полукруглые и квадратные скобки, слеш и т. д.) (примеры 3–5);
- (3) использование шрифтового оформления в тексте статей (полужирный шрифт, курсив, прописные буквы, в частности для обозначения падежей в немецком языке) (см. курсив в примере 6).

Эти правила служат основой создания алгоритма и программы структуризации текста наследуемых словарных статей до их загрузки в БД. Первое правило представлено таблицей 1 в работе [4]. Оно задает размещение компонентов, зон и полей в статье. Размещение обязательных *нечисловых* компонентов статьи иллюстрируется примером 1, а компонентов, зон и полей в статье с идиомой, которая имеет два пронумерованных значения, — примером 2.

Второе правило определяет формат дат (см. название источника и дату в европейском формате в примере 1 внутри круглых скобок), а также позиции использования трех символов шрифта Wingdings (☞ — для каждого из примеров в словарной статье; ☞ — для зоны вариантов; ☞ — для зоны комментариев) и спецсимволов. Пример 1 включает только один символ ☞. Рассмотрим примеры использования остальных двух символов начала зон вариантов и комментариев

Пример 2 Описание двух значений идиомы в словарной статье

einen Tritt geben

1. дать пинка (*кому-л.*); попытаться повлиять (*на кого-л.*)

☞ Bevor er sich an den Schreibtisch setze, habe er jeden Tag “Mulmigkeit und Bangigkeit” zu überwinden. Manchmal müsse er *sich* regelrecht *einen Tritt geben*. (Nach: Nürnberger Zeitung, 23.01.2008)

Прежде чем сесть за письменный стол, он вынужден каждый день бороться с «апатией и страхом». Иногда ему приходится просто *заставлять себя*.

2. дать пинка (*кому-л.*); выставить [вышвырнуть] на улицу (*кого-л. — уволить*)

☞ Auf keinen Fall haben die Politiker ihre Pflicht erfüllt. In der freien Wirtschaft *gibt* man solch unfähigen Leuten *einen Tritt*. (Nach: Nürnberger Zeitung, 10.01.2011)

Политики ни в коей мере не выполнили свою задачу. В бизнесе таким бестолковым людям *дают пинка*.

Пример 3 Использование символа начала зоны вариантов

sich den Bauch vollschlagen

набить желудок, наесться до отвала; набить себе брюхо (*снижен.*)

☞ **sich den Magen vollschlagen**

в словарных статьях. Зона вариантов является опциональной, следует за зоной перевода и вводится символом ☞ (пример 3).

В приведенном варианте идиомы используется синоним Magen для слова Bauch (желудок). Зона комментариев также является опциональной, может располагаться в любой части словарной статьи (в зависимости от характера сообщаемых сведений) и вводится символом ☞ (пример 4).

Кроме трех символов шрифта Wingdings в словарной статье используются спецсимволы (полукруглые и квадратные скобки, слеш и т. д.). Спецсимволы могут встречаться в любой зоне словарной статьи и использоваться с разной целью. В частности, в примере 4 в русском переводе полукруглые скобки указывают альтернативные валентности *что-л.*, *кого-л.*, разделенные запятой. При этом разные морфосинтаксические способы заполнения одной и той же валентности указываются через слеш (пример 5). Квадратными скобками могут быть выделены лексические варианты (в примере 1 в идиоме *nach mir [uns]*

Пример 4 Использование символа начала зоны комментария

es abgesehen haben

(*что-л.*) приглянулось (*кому-л.*), (*кто-л.*) положил глаз (*на что-л.*, *кого-л.*)

☞ Идиома преимущественно употребляется для обозначения преступных намерений субъекта.

Пример 5 Использование слеша при описании валентностей

grünes Licht bekommen

получить разрешение (*на что-л.*), получить зеленый свет (*на что-л. / для чего-л.*)

die Sintflut и ее переводе *после меня [нас] хоть потоп* в скобках дан вариант использования идиомы во множественном числе).

Третье правило говорит о применении шрифтового оформления в тексте статей. Элементы оформления могут присутствовать в любой зоне словарной статьи. Основным элементом оформления служит выделение полужирным шрифтом или курсивом. Например, идиома всегда выделяется полужирным шрифтом (см. идиому **grünes Licht bekommen** в примере 5). Курсивом выделяются валентности идиомы (см. *на что-л. / для чего-л.* в примере 5), поясняющий или уточняющий комментарий (см. пример 6) и для выделения идиомы внутри примера ее использования и перевода (см. в примере 1 «*после меня хоть потоп*» и «*после нас хоть потоп*»).

Пример 6 Поясняющий комментарий в переводе идиомы

gleich null sein

быть равным нулю (*напр. о результате*), быть ничтожным

Прописные буквы в словарных статьях могут встречать в разных зонах, например для выделения вокабулы или указания падежей немецкого языка.

2.2 Второй принцип

После завершения процесса структурирования словарных статей и их загрузки в БД открывается возможность дополнения словарных статей новыми полями и их значениями, отсутствующими в наследуемых ресурсах. Основная цель их дополнения состоит в многоаспектном рубрицировании статей.

Рубрицирование словарных статей планируется реализовать как минимум в двух аспектах. Во-первых, в словарную статью добавляется поле рубрик смысловой категории согласно «Тезаурусу русских идиом» [5]. Этот тезаурус включает около 6000 идиом русского языка, сгруппированных по тематическим категориям, которые образуют многоуровневую иерархию. Верхний уровень этой иерархии включает 87 категорий: «1. Время», «2. Пространство», «3. Движение», . . . , «87. Вещество». Согласно этой тематической категоризации идиома **sich den Bauch vollschlagen** из примера 3 попадает в категорию верхнего уровня «64. Еда. Пища» и получает рубрику «64.3. Много еды, много есть, чревоугодие». А идиома **es abgesehen haben** из примера 4 (когда речь идет не о преступных намерениях) попадает в две категории верхнего уровня:

- «25. Эмоциональные/психические состояния, чувства» и получает первую рубрику «25.4.1. Заинтересованность»;
- «66. Любовь, брак, сексуальные отношения» и получает вторую рубрику «66.2. Флирт, ухаживания».

Добавление рубрик тематической категории дает возможность группировать идиомы согласно рубрике любого уровня, включая любую из 87 категорий верхнего уровня.

Во-вторых, в словарную статью добавляются поля для рубрик классификационной схемы семантических, грамматических, сочетаемостных и стилистических признаков в интересах лингвистической разметки статей, описанию которой планируется посвятить отдельную работу. Отметим, что тематическая категоризация не совпадает с семантическими признаками в том случае, если в словаре одновременно используются две разные семантические классификации.

3 Сопоставительный анализ

Существуют лексикографические системы формирования словарей, по ряду функций аналогичные БД фразеологического словаря. Для сопоставительного анализа в статье используются труды конференции по Dictionary Writing Systems (DWS), в которых описываются наиболее известные системы, в том числе те, сравнение с которыми будет проведено ниже, а именно: TschwaneLex (TLex), Lexonomy, Lexique Pro, Mataruna [6].

Сопоставим функционал БД фразеологического словаря с функциональными возможностями вышеупомянутых систем с точки зрения импорта словарных статей. Почти во всех этих системах реализован первый принцип. Они используют основополагающие правила, которых придерживаются лексикографы: соблюдение порядка размещения компонентов, единообразное использование шрифтового оформления и спецсимволов. Отметим, что не все DWS обязательно используют символы для разграничения зон. Такая функция присутствует только в TLex и Lexique Pro. Кроме того, в них реализована функция пользовательской настройки единообразного шрифтового оформления текста статей, в том числе с использованием цвета.

В электронных словарях статьи структурированы согласно схеме, задаваемой лексикографами. Однако уже после начала процесса создания словаря у них может возникнуть потребность изменить эту схему. Функция ее изменения, включая дополнение словарных статей новыми структурными элементами, присутствует, например, в системе TLex (при помощи функции Document Type Definition [7]). Есть эта функция и в системе Lexonomy — при желании лексикограф может использовать структурную схему по умолчанию, дополнять или изменять ее в ходе работы или же создать полностью новую структуру [8]. Новый структурный элемент может быть добавлен на любом уровне иерархии схемы —

возможно создание новых компонентов, зон внутри компонентов, а также полей внутри зон.

Ряд ключевых отличий системы БД фразеологического словаря от рассматриваемых систем обнаруживается при сравнении их по второму принципу — дополнению словарных статей новыми полями и их значениями, отсутствующими в наследуемых ресурсах. Рассмотрим функцию добавления рубрик систем классификаций в словарные статьи. Среди рассматриваемых систем такой функцией обладает Lexique Pro. Она позволяет пользователю создать собственную классификацию тематических рубрик, а затем распределять словарные статьи по ним. Категории, обозначаемые рубриками, могут быть многоуровневыми. Существует возможность по умолчанию использовать классификацию и ее категории, взятые из программы Dictionary Development Program, разработанной Рональдом Моэ [9]. Эта классификация обладает многоуровневой структурой, и почти каждая категория имеет вложенные рубрики (например, 3. Language and thought → 3.2. Mind → 3.2.5. Opinion → 3.2.5.1. Believe).

В первой версии фразеологического словаря, создаваемой с помощью проектируемой БД, предусмотрено применение рубрик четырех систем классификации (семантическая, стилистическая, грамматическая и сочетаемостная), включая семантическую классификацию Баранова–Добровольского, используемую в «Тезаурусе русских идиом» [5]. Другими словами, используются рубрики только одной семантической классификации.

Отметим, что система TLex не обладает функцией распределения статей по тематическим категориям. Однако существует возможность эмулировать данную функцию, используя систему тегов. Правда, такой способ обеспечивает функционал, который будет значительно редуцирован по сравнению с Lexique Pro или БД фразеологического словаря. Причина в том, что эта система тегов позволяет создавать лишь рубрики первого уровня, т. е. в форме одноуровневых списков. Кроме того, в этом случае можно задать только одну классификацию в отличие от БД фразеологического словаря, которая сможет поддерживать одновременно несколько классификаций.

Таким образом, в результате сопоставительного анализа рассмотренных лексикографических систем с БД фразеологического словаря можно сделать вывод о том, что ни одна из вышеупомянутых систем по отдельности не обладает ее функционалом в части импорта наследуемых лексикографических ресурсов. Для наглядного представления результатов сопоставительного анализа приведем таблицу с тремя функциями:

- (1) автоматизированный импорт наследуемых ресурсов;
- (2) добавление элементов в схему структуризации словарных статей;
- (3) создание и использование систем классификации.

Рассмотрев таблицу, можно сделать вывод о том, что на настоящий момент времени наиболее полный функционал в части импорта словарных статей предусмотрен в БД фразеологического словаря.

Сопоставление лексикографических систем

Функции DWS	TLex	Lexonomy	Lexique Pro	Matapuna
Импорт	✓	—	—	—
Добавление элементов	✓	✓	✓	—
Системы классификации	—	—	✓	—

4 Заключение

Наследуемый лексикографический ресурс БД фразеологического словаря — это около 2000 словарных статей в формате текстовых документов. Для разметки статей использованы функции форматирования редактора Word — выделение полужирным шрифтом, курсивом и т. д.

При использовании средств и систем информатики при формировании электронного словаря и работы с ним необходимо структурировать наследуемые лексикографические ресурсы и пополнять словарные статьи новыми полями, что существенно расширяет спектр пользовательских функций при работе со словарем, включая семантический, грамматический и другие виды поиска.

Для формирования БД фразеологического словаря и обеспечения пользовательских функций при работе с ним используется система управления базами данных SQL Server. База данных содержит наследуемую и новую информацию, добавляемую лексикографами. Функционал БД фразеологического словаря объединяет в части импорта и обогащения словарных статей возможности наиболее известных лексикографических систем TLex, Lexonomy и Lexique Pro.

Литература

1. *De Schryver G.-M.* Lexicographers' dreams in the electronic-dictionary age // *Int. J. Lexicogr.*, 2003. Vol. 16. No. 2. P. 143–199.
2. *Tarp S.* Lexicography in the Borderland between knowledge and non-knowledge: General lexicographical theory with particular focus on learner's lexicography. — Tübingen: Max Niemeyer Verlag, 2008. 308 p.
3. *Müller-Spitzer C.* Textual structures in electronic dictionaries compared with printed dictionaries: A short general survey // *Dictionaries: An international encyclopedia of lexicography. Supplementary Volume: Recent developments with focus on electronic and computational lexicography.* — Berlin: De Gruyter Mouton, 2013. P. 367–381.
4. *Вакуленко В. В., Гончаров А. А., Дурново А. А., Зацман И. М.* Задачи базы данных фразеологического словаря и стадии ее проектирования // *Системы и средства информатики*, 2020. Т. 30. № 2. С. 113–123.
5. *Тезаурус русских идиом: семантические группы и контексты* / Под ред. А. Н. Баранова, Д. О. Добровольского. — М.: Лексрус, 2018. 888 с.
6. *Suhardinato T., Dinakaramani A.* Building a collaborative workspace for lexicography works in Indonesia // *Electronic Lexicography in the 21st Century Conference Pro-*

- ceedings / Eds. I. Kosem and C. Tiberius. — Brno: Lexical Computing CZ s. r. o., 2017. P. 299–308.
7. TLex Suite User Guide: Professional lexicography, terminology and corpus query software. Version 14.0.27. <https://www.tshwanedje.com/docs/TLex%20Suite%20User%20Guide.pdf>.
 8. *Měchura M. B.* Introducing lexonomy: An open-source dictionary writing and publishing system // *Electronic Lexicography in the 21st Century: eLex 2017 Conference Proceedings* / I. Kosem and C. Tiberius. — Brno: Lexical Computing CZ s. r. o., 2017. P. 662–679.
 9. *Moe R.* Dictionary development program // *SIL Forum for Language Fieldwork, 2007. Series Issue 2007-003. Entry Number 7770. 12 p.* <https://www.sil.org/resources/archives/7770>.

Поступила в редакцию 15.03.21

INHERITABLE LEXICOGRAPHIC RESOURCES OF THE PHRASEOLOGICAL DICTIONARY DATABASE

V. V. Vakulenko and I. M. Zatsman

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

Abstract: The paper presents an overview of the principles of importing pre-existing dictionary articles into the database of the phraseological dictionary, i. e., using inheritable lexicographic resources. The import of dictionary articles into the database is based on two key principles: structuring the text of the articles and adding new fields and their values into the articles. Each article contains an idiom in German, one or several translations into Russian, their contexts, and a range of other components. The goal of the paper is to describe the principles of using and supplementing inheritable lexicographic resources during the process of forming the database of the phraseological dictionary of German idioms as well as comparing the features of the database with capabilities of similar lexicographic information systems of dictionary creation from the standpoint of the two aforementioned principles of import.

Keywords: import principles; inheritable lexicographic resources; comparative analysis; phraseological dictionary database

DOI: 10.14357/08696527210212

References

1. De Schryver, G.-M. 2003. Lexicographers’ dreams in the electronic-dictionary age. *Int. J. Lexicogr.* 16(2):143–199.

2. Tarp, S. 2008. Lexicography in the borderland between knowledge and non-knowledge. *General lexicographical theory with particular focus on learner's lexicography*. Tübingen: Max Niemeyer Verlag. 308 p.
3. Müller-Spitzer, C. 2013. Textual structures in electronic dictionaries compared with printed dictionaries. A short general survey. *Dictionaries. An international encyclopedia of lexicography. Supplementary Volume: Recent developments with focus on electronic and computational lexicography*. Berlin: De Gruyter Mouton. 367–381.
4. Vakulenko, V. V., A. A. Goncharov, A. A. Durnovo, and I. M. Zatsman. 2020. Zadachi bazy dannyykh frazeologicheskogo slovarya i stadii ee proektirovaniya [Tasks of the phraseological dictionary database and stages of its design]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(2):113–123.
5. Baranov, A. N., and D. O. Dobrovol'skiy, eds. 2018. *Tezaurus russkikh idiom. Semanticheskie gruppy i konteksty* [Thesaurus of Russian idioms. Semantic groups and contexts]. Moscow: Lekrus. 888 p.
6. Suhardijanto, T., and A. Dinakaramani. 2017. Building a collaborative workspace for lexicography works in Indonesia. *Electronic Lexicography in the 21st Century Conference Proceedings*. Eds. I. Kosem and C. Tiberius. Brno: Lexical Computing CZ s. r. o. 299–308.
7. TLex Suite User Guide. Available at: <https://www.tshwanedje.com/docs/TLex%20Suite%20User%20Guide.pdf> (accessed March 29, 2021).
8. Měchura, M. B. 2017. Introducing Lexonomy: An open-source dictionary writing and publishing system. *Lexicography in the 21st Century Conference Proceedings*. Eds. I. Kosem and C. Tiberius. Brno: Lexical Computing CZ s. r. o. 662–679.
9. Moe, R. 2007. Dictionary development program. SIL Forum for Language Fieldwork 2007-003:7770. Available at: <https://www.sil.org/resources/archives/7770> (accessed March 29, 2021).

Received March 15, 2021

Contributors

Vakulenko Vasily V. (b. 1995) — engineer-researcher, 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; vvak@pm.me

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

ИНДИКАТОРНАЯ ОЦЕНКА НЕСТАБИЛЬНОСТИ НЕЙРОННОГО МАШИННОГО ПЕРЕВОДА

А. Ю. Егорова¹, И. М. Зацман², М. Г. Кружков³, В. А. Нуриев⁴

Аннотация: Представлены данные, полученные в результате наблюдения за работой системы нейронного машинного перевода (НМП), а также анализа допущенных в переводе ошибок. В ходе эксперимента получена количественная оценка нестабильности НМП с помощью индикаторов. Экспериментальным материалом послужили 250 русскоязычных текстовых фрагментов, для каждого из которых ежемесячно в течение одного года фиксировался как перевод на французский язык, выполненный с помощью системы НМП Google Translate, так и допущенные в нем ошибки. Фиксация ошибок проведена в процессе аннотирования переводов с помощью надкорпусной базы данных (НБД). В результате ежемесячной обработки переводов была получена серия из 12 аннотаций для каждого из 250 текстовых фрагментов. Аннотирование переводов позволило не только зафиксировать допущенные в переводе ошибки в случае их наличия, но и определить категории нестабильности НМП, указывающие на изменения качества перевода или на отсутствие ошибок. Цель статьи состоит в сопоставительном анализе результатов работы Google Translate с учетом временного параметра.

Ключевые слова: нейронный машинный перевод (НМП); нестабильность машинного перевода; надкорпусная база данных; индикаторная оценка; лингвистическое аннотирование; категории, характеризующие нестабильность НМП

DOI: 10.14357/08696527210213

1 Введение

В настоящее время онлайн-системы машинного перевода (МП) находят применение во многих сферах человеческой жизни и развиваются быстрыми темпами. Одновременно ведутся исследования по экспертному (с привлечением человека-

¹Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, ann.shurova@gmail.com

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

³Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, magnit75@yandex.ru

⁴Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, nurieff.v@gmail.com

эксперта) и автоматическому сопоставительному анализу работы разных систем. Последний тип анализа связан с разработкой и использованием метрик автоматической оценки качества МП. Автоматическая оценка измеряет уровень соответствия МП референтному — выполненному человеком-профессионалом — переводу (РП). Чтобы установить уровень соответствия МП и РП, применяются критерии точности (процент переведенного без ошибок) и полноты (доля действительно переведенного). Самой распространенной метрикой автоматической оценки качества МП служит BLEU (*англ.* Bilingual Evaluation Understudy). Определяя соответствие МП и РП, BLEU подсчитывает n -граммную точность и назначает штраф за слишком короткий перевод. N -граммная точность при этом представляет собой отношение последовательностей из n слов, совпадающих в МП и РП, к общему числу последовательностей из n слов в МП. BLEU основывается на статистическом подходе, не используя знания о контактирующих при переводе языках. В отличие от нее, другая наиболее востребованная метрика, METEOR (*англ.* Metric for Evaluation of Translation with Explicit Ordering), предполагает интеграцию лингвистического знания. Так, она учитывает не только полные n -граммные совпадения в МП и РП, но и использует корни слов вместо точных словоформ, найденных в текстах, а также семантические данные (подробнее об автоматизированных способах оценки качества МП см. [1–4]).

В процессе экспертного сопоставительного анализа используется, как правило, авторская классификация тех ошибок, которые регистрируются по результатам перевода. Например, в работе [5] был проведен экспертный анализ функционирования четырех онлайн-систем машинного перевода: Bing Translator, Google Translate, Systran и Yandex.Translate. Экспериментальный массив включал 300 предложений на английском языке по медицинской тематике. Они были переведены на хорватский язык с помощью этих четырех систем. В полученных переводах регистрировались девять видов ошибок, включая неправильный перевод многозначного слова. Например, английское слово «medicine» может означать медицина и лекарство. Также регистрировались непереведенные слова и ошибки при переводе идиом. Меньше всего ошибок (по всем девяти их видам) сделали Google Translate (89) и Yandex.Translate (92), больше всего — система Systran (213). У Bing Translator было зарегистрировано 138 ошибок. Все четыре системы допустили наибольшее число ошибок при переводе многозначных слов.

Цель данной статьи состоит в экспертном сопоставительном анализе результатов работы одной и той же онлайн-системы НМП Google Translate в разные моменты времени на одном и том же экспериментальном массиве предложений. Такой подход дает возможность определить степень стабильности в работе системы на основе ранее разработанной методики регистрации и анализа ошибок перевода [6]. В методике использовалась авторская классификация ошибок, которые регистрируются и анализируются в процессе аннотирования оригинала и перевода.

2 Описание эксперимента

Материалом для эксперимента послужили 250 русскоязычных предложений, каждое из которых содержит двухкомпонентный коннектор¹ (например, *как (расстояние) так и*). Источником всех предложений был Национальный корпус русского языка (НКРЯ) [8]. Каждое предложение в течение года было переведено на французский язык и проаннотировано вместе с его МП 12 раз.

Для формирования, хранения и поиска аннотаций применялась НБД коннекторов [9–13]. Каждая аннотация (см. пример с коннектором *как (расстояние) так и* в табл. 1) состоит из следующих частей: предложение или его фрагмент с коннектором на русском языке, его перевод на французский, выполненный с помощью системы НМП Google Translate, и структурированное описание ошибок в случае их наличия.

Каждый месяц на протяжении года (с марта 2019 г. по февраль 2020 г.) в НБД коннекторов происходило формирование 250 аннотаций с МП отобранных предложений или их фрагментов. В результате был получен массив экспериментальных данных, в котором для каждого из 250 предложений была зафиксирована серия из 12 версий МП. Всего в НБД было сформировано 3000 аннотаций.

Для обработки всех сформированных аннотаций была применена методика регистрации и анализа ошибок НМП [6] и авторская классификация количественных показателей (индикаторов) ошибок (далее — индикаторная оценка ошибок, ИО). Классификация индикаторов и методика были разработаны в три этапа. На первом этапе, описанном в [16], была проведена категоризация нестабильности НМП на основе анализа ошибок перевода, которые возникали или исчезали в сериях МП при НМП.

В результате были выделены шесть категорий. Пять из них характеризуют нестабильность НМП:

- (1) повышение качества НМП;
- (2) снижение качества НМП;
- (3) колебание качества НМП;
- (4) изменение набора ошибок в НМП без динамики его качества;
- (5) изменение НМП без динамики его качества.

Шестая категория объединяет переводы, которые оставались неизменными в пределах серии (описание категорий представлено в [16]).

На втором этапе была разработана и описана методика регистрации и анализа ошибок НМП [6]. На третьем этапе исследования [17] была проведена предварительная оценка результатов работы системы НМП с применением разработанной классификации индикаторов (на материале 1800 аннотаций).

¹Коннектор — языковая единица, функция которой состоит в выражении логико-семантического отношения, существующего между соединенными с ее помощью частями текстового фрагмента [7, с. 17].

Таблица 1 Пример аннотации

№ в НБД	Контекст РР в оригинале ¹	РР в оригинале и ее рубрики	Контекст РР в переводе ²	РР в переводе и ее рубрики	Признаки аннотации	Дата и время записи в НБД
28323	Как 20–30 лет назад, так и теперь, перед смертью, меня интересует одна только наука [А. П. Чехов. Скучная история (1889)]	как так и (аналогия) (CNT p CNT q) (CNT) (Дистант)	Comme il y a 20–30 ans, alors maintenant, avant la mort, je ne m'intéresse qu'à la science [перевод выполнен системой Google Translate]	comme alors (ErrorTotal CNT)	MT	03.03.2019 17:27

¹Речевая реализация в оригинале — форма коннектора в конкретном высказывании [14].

²Речевая реализация в переводе — функционально эквивалентный фрагмент перевода, передающий смысловое наполнение РР в оригинале. Понятие «функционально эквивалентный фрагмент» предложено в [15].

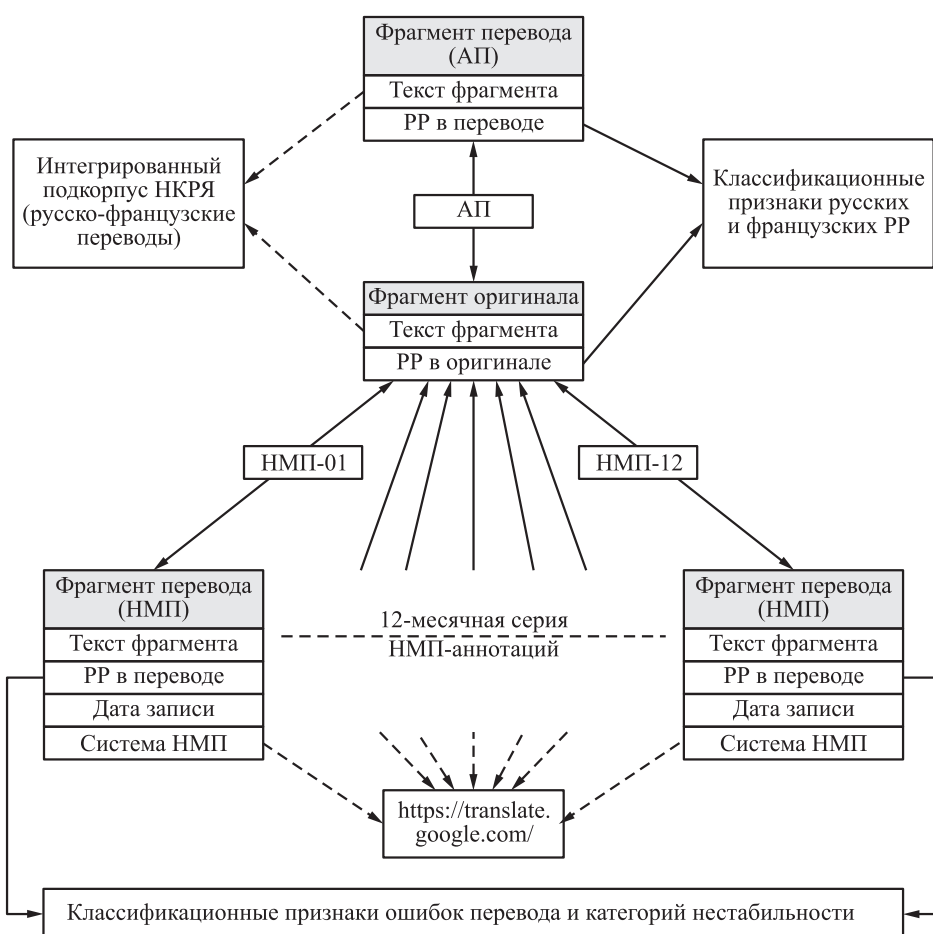
Примечания.

1. Рубрика (подробнее о рубриках см. в [11]) «аналогия» говорит о том, что коннектор **как||так и** выражает логико-семантическое отношение аналогии.
2. Рубрика «CNT p CNT q» говорит о том, что элементы двухкомпонентного коннектора находятся в каждом из соединяемых компонентов текста *p* («20–30 лет назад») и *q* («теперь, перед смертью, меня интересует одна только наука»).
3. Рубрика «CNT» говорит о том, что аннотация сформирована для всего коннектора, а не для отдельно составляющих его блоков или элементов.
4. Рубрика «Дистант» говорит о том, что составные части коннектора «как» и «так и» разделены текстом «20–30 лет назад».
5. Рубрика «ErrorTotalCNT» говорит о том, что коннектор **как||так и** переведен языковой единицей **comme||alors** (несуществующей во французском языке).
6. Рубрика «MT» проставляется в аннотации программно при использовании машинного переводчика.

3 Адаптация надкорпусной базы данных для индикаторной оценки

Чтобы фиксировать в НБД коннекторов информацию о переводах текстовых фрагментов, содержащих русские коннекторы, на французский язык с использованием машинных переводчиков, эта НБД была дополнена рядом функций, позволяющих сохранять информацию о переводах, получаемых из внешних источников (в данном случае — из систем НМП). Изначально НБД коннекторов позволяла формировать аннотации только для переводов, извлекаемых из интегрированного в НБД многоязычного подкорпуса НКРЯ [8], который содержит авторские переводы (АП), т. е. переводы, выполненные профессиональными переводчиками.

Аннотации для НМП строятся по тому же принципу, что и аннотации для АП (см., например, [18]), но при этом они обладают следующими особенностями. Во-первых, вместо ссылки на соответствующие фрагменты корпуса они включают в себя ссылки на систему НМП, с помощью которой был выполнен приведенный перевод, а также сохраняют дату перевода. Во-вторых, в рамках описываемого эксперимента для каждого фрагмента оригинала в течение года формировалось по серии из 12 НМП-аннотаций, в результате чего в НБД создаются информационные объекты нового типа (темпоральные серии НМП-аннотаций), позволяющие отслеживать динамику изменений в МП во времени.



Структурная схема НБД коннекторов с поддержкой аннотирования серий НМП из внешних источников

Наконец, фасетная классификация для НМП-аннотаций проводится по другим основаниям, отличным от оснований, используемых при классификации АП-аннотаций. При этом главным образом используется два новых фасета: фасет для классификации ошибок перевода, позволяющий фиксировать наличие или отсутствие различных видов ошибок в конкретных МП, и фасет категорий нестабильности, позволяющий фиксировать экспертную оценку динамики качества НМП для каждой серии НМП-аннотаций (подробное описание этих фасетов дано в работах [6, 16]). Структурная схема центральной части НБД коннекторов с учетом вышеописанных изменений приведена на рисунке.

Данная схема позволяет извлекать из НБД коннекторов сведения о корреляциях между различными аспектами оригинальных и переводных текстов (такими как речевые реализации (РР) и классификационные признаки НМП, АП и оригинальных текстов, распределение НМП по месяцам, признаки переводных соответствий и т. д.). По запросу пользователей администраторы НБД формируют SQL-запросы с агрегирующими функциями, генерирующие количественные и процентные данные о таких корреляциях, которые затем интерпретируются как индикаторы динамики МП.

Так, в [17] приводились предварительные данные о корреляциях между категориями нестабильности НМП и ошибками, наблюдаемыми в МП, относящихся к этим категориям. Также были представлены данные о распределении различных видов ошибок в НМП по месяцам.

4 Результаты индикаторной оценки

В рамках эксперимента по мониторингу работы системы НМП были получены числовые индикаторы, характеризующие соотношение категорий нестабильности НМП, ошибок в МП и случаев их отсутствия. В табл. 2 представлены данные для случаев, когда ошибки в МП отсутствуют (строка 1), а также данные о частотности 19 видов ошибок МП¹ (см. строки 2–20). Частотность вычислялась от общего числа ошибок (4494), зарегистрированных в 3000 аннотациях, при этом в одной аннотации могло быть зафиксировано несколько разных ошибок. В НБД коннекторов была добавлена возможность проставлять для одной аннотации несколько ошибок одного вида, что позволило уточнить предварительные данные эксперимента [17].

Прежде всего, отметим, что для 645 аннотаций из 3000 был проставлен признак NoError. Это значит, что 21,5% всех МП представляют собой варианты перевода без ошибок, во всех остальных аннотациях были зафиксированы одна или более ошибок.

Самой частотной стала ошибка ErrorSemant (43,01% от 4494, т. е. от суммы всех проставленных кодов ошибок и случаев их отсутствия, когда проставляется код NoError). Это лексическая ошибка, которая вызывает искажение смысла во

¹Подробнее о классификации ошибок см. в [6, 16].

Таблица 2 Индикаторы нестабильности НМП, ошибок и случаев их отсутствия

№	Код ошибки	1	2	3	4	5	6	Сумма по 6 категориям, характеризующим нестабильность	Значения индикаторов ошибок, %
		Повышение качества НМП	Снижение качества НМП	Колебание качества НМП	Изменение набора ошибок в НМП без динамики его качества	Изменение НМП без динамики его качества	НМП без изменений		
1	NoError	116	81	148	0	252	48	645	14,35
2	ErrorSemant	601	169	657	325	169	12	1933	43,01
3	ErrorSyntax	100	34	158	102	24	0	418	9,30
4	ErrorSyntaxPostCNT	109	35	92	72	48	0	356	7,92
5	ErrorCNT	92	7	46	84	29	0	258	5,63
6	ErrorTotalCNT	58	8	50	65	24	0	205	4,67
7	Pleonasm	24	12	23	23	23	0	105	2,34
8	ErrorPunct	21	15	33	27	0	0	96	2,14
9	ErrorPart1CNT	22	0	31	21	12	0	86	1,91
10	Lacuna	20	0	25	35	0	0	80	1,78
11	ErrorMorphCNT	16	0	10	29	12	0	67	1,49
12	ErrorPart2CNT	22	3	16	14	12	0	67	1,49
13	ErrorMorphPostCNT	5	5	31	23	0	0	64	1,42
14	ErrorMorph	5	3	24	6	0	0	38	0,85
15	Latin	15	0	2	12	0	0	29	0,65
16	TrPart2CNT	2	0	4	0	12	0	18	0,40
17	AgramTotal	5	0	0	12	0	0	17	0,38
18	TrPart1CNT	0	0	5	7	0	0	12	0,27
19	Cyrillic	0	0	0	0	0	0	0	0,00
20	ErrorOrthCNT	0	0	0	0	0	0	0	0,00
Сумма по видам ошибок		1233	372	1355	857	617	60	4494	
Значения индикаторов для каждой категории, %		27,44	8,28	30,15	19,07	13,73	1,33		100,00

фрагменте перевода. В 3000 аннотаций она была проставлена 1933 раза, причем нередко в одном МП фиксировались несколько ошибок вида ErrorSemant.

Далее по уменьшению частотности следуют синтаксические ошибки во фрагменте, не вводимом коннектором (ErrorSyntax — 9,3%) и вводимом коннектором (ErrorSyntaxPostCNT — 7,92%). Синтаксические ошибки встречаются в МП гораздо реже, чем ErrorSemant: если объединить ErrorSyntax и ErrorSyntaxPostCNT вместе, их число будет почти в 2,5 раза меньше, чем число ErrorSemant.

Затем следуют морфологические ошибки (ErrorMorphPostCNT для фрагмента с коннектором и ErrorMorph для фрагмента, не вводимого коннектором),

Таблица 3 Распределение аннотаций по категориям, характеризующим нестабильность НМП

№	Название рубрики, характеризующей нестабильность НМП	Число аннотаций	Доля от общего числа аннотаций (3000)
1	Повышение качества НМП	828	27,6%
2	Снижение качества НМП	288	9,6%
3	Колебание качества НМП	924	30,8%
4	Изменение набора ошибок в НМП без динамики его качества	408	13,6%
5	Изменение НМП без динамики его качества	492	16,4%
6	НМП без изменений	60	2,0%

частотность которых составила 1,42% и 0,85% соответственно. Значительная разница между числом синтаксических и морфологических ошибок подтверждает тенденцию, выявленную на этапе анализа предварительных данных по 1800 аннотациям, которая была описана в [17].

Что касается ошибок, затрагивающих непосредственно перевод коннекторов, самыми распространенными стали ErrorCNT (в переводе используется семантически неверный коннектор) и ErrorTotalCNT (система НМП переводит коннектор несуществующей языковой единицей). Их частотность равна 5,63% и 4,67% соответственно.

По завершении исследования НМП нулевыми остались ячейки для ошибок Cyrillic (использование в переводе слова кириллицей) и ErrorOrthCNT (орфографическая ошибка в форме коннектора). Как уже было отмечено в [17], эти показатели могут свидетельствовать о тенденции к повышению качества НМП, учитывая тот факт, что в предыдущих исследованиях [19] эти значения не были нулевыми.

Анализируя данные по первым пяти категориям, характеризующим нестабильность НМП, отметим следующее. Третьей категории «Колебание качества НМП» и первой категории «Повышение качества НМП» соответствует наибольшее число ошибок (30,15% и 27,44% соответственно), что коррелирует с тенденцией распределения аннотаций по категориям, характеризующим нестабильность НМП (табл. 3).

5 Количественные показатели нестабильности нейронного машинного перевода

Эксперимент по наблюдению за работой НМП также позволил получить количественные данные распределения аннотаций по категориям, характеризующим нестабильность НМП (их примеры см. в [16, 17]).

В табл. 3 представлены данные по всем 3000 аннотаций. Следует отметить, что здесь процентное соотношение аннотаций несколько отличается от промежуточных данных по 1800 аннотациям, описанных в [17]. Так, по данным в табл. 3 наибольшее значение индикатора (30,8% аннотаций) у третьей категории, характеризующей колебание качества НМП. На втором месте по числу аннотаций — первая категория, характеризующая повышение качества перевода (27,6% всех аннотаций). Важно отметить, что по промежуточным данным самой частотной была первая категория (на промежуточном этапе было зафиксировано 30,6% таких аннотаций). Такое изменение соотношения категорий, возможно, говорит о том, что по мере обучения качество НМП не всегда изменяется только в сторону повышения или снижения. Оно колеблется, т. е. число допущенных в переводе ошибок растет или уменьшается, при этом их состав может меняться.

Затем следуют пятая категория, характеризующая изменение НМП без динамики его качества (16,4%), и четвертая категория, характеризующая изменение набора ошибок в переводе без динамики его качества (13,6%). Примечательно, что число аннотаций третьей и первой категорий превышает число аннотаций пятой и четвертой категорий почти в 2 раза. При этом соотношение пятой и четвертой категорий также немного изменилось по сравнению с промежуточными данными, когда к четвертой категории относилось большее число аннотаций, чем к пятой.

Наименьшее число аннотаций приходится на вторую категорию, характеризующую снижение качества перевода (9,6%). Стабильными на протяжении эксперимента остались 2% переводов (НМП без изменений).

6 Заключение

В статье описаны результаты эксперимента по мониторингу НМП посредством многократного фиксирования переводов, сделанных с помощью системы Google Translate. В ходе эксперимента осуществлялась индикаторная оценка работы НМП, в результате чего были получены количественные показатели соотношения категорий нестабильности НМП, ошибок МП и случаев их отсутствия. Самой частотной для всех шести категорий, характеризующих нестабильность НМП, стала ошибка ErrorSemant (43,01%). Это может свидетельствовать о том, что основные трудности в работе системы НМП Google Translate возникают не на грамматическом, а на лексико-семантическом уровне.

Также в результате проведенного эксперимента были получены данные о распределении аннотаций по категориям, характеризующим нестабильность НМП. Наибольшее число аннотаций относится к третьей категории — «Колебание качества НМП» (30,8%). Такое распределение аннотаций может указывать на тот факт, что со временем эффективность работы системы НМП не всегда изменяется только в одном направлении: качество перевода может колебаться, а число и состав допущенных в переводе ошибок — меняться. Результаты эксперимента

и изложенные выводы подтверждают актуальность дальнейших наблюдений за работой НМП и проведения индикаторной оценки результатов МП.

Литература

1. *Banerjee S., Lavie A.* METEOR: An automatic metric for MT evaluation with improved correlation with human judgments // 43rd Annual Meeting of the Association of Computational Linguistics and the Workshop on Intrinsic and Extrinsic Evaluation Measures for MT and/or Summarization Proceedings. — Ann Arbor, MI, USA: Association of Computational Linguistics, 2005. P. 65–72.
2. Translation quality assessment. From principles to practice / Eds. J. Moorkens, Sh. Castilho, F. Gaspari, S. Doherty. — Machine translation: Technologies and applications ser. — Cham, Switzerland: Springer International Publishing, 2018. Vol. 1. 299 p.
3. *Specia L., Scarton C., Paetzold G. H.* Quality estimation for machine translation. — San Rafael, CA, USA: Morgan & Claypool, 2018. 162 p.
4. *Koehn Ph.* Neural machine translation. — New York, NY, USA: Cambridge University Press, 2020. 394 p.
5. *Mikulić A.* Ljudska evaluacija sustava za neuralno strojno prevođenje (Završni rad). 2020. <https://urn.nsk.hr/urn:nbn:hr:131:516896>.
6. *Егорова А. Ю., Зацман И. М., Кружков М. Г., Нуриев В. А.* Методика темпоральной оценки нестабильности машинного перевода // Системы и средства информатики, 2020. Т. 30. № 3. С. 67–80.
7. *Инькова-Манзотти О. Ю.* Коннекторы противопоставления во французском и русском языках. Сопоставительное исследование. — М.: Информэлектро, 2001. 429 с.
8. Национальный корпус русского языка. <http://www.ruscorpora.ru>.
9. *Зализняк А. А., Зацман И. М., Инькова О. Ю., Кружков М. Г.* Надкорпусные базы данных как лингвистический ресурс // Корпусная лингвистика: Тр. VII Международ. конф. — СПб.: СПбГУ, 2015. С. 211–218.
10. *Дурново А. А., Зацман И. М., Лоццлова Е. Ю.* Кросслингвистическая база данных для аннотирования логико-семантических отношений в тексте // Системы и средства информатики, 2016. Т. 26. № 4. С. 124–137.
11. *Зализняк А. А., Зацман И. М., Инькова О. Ю.* Надкорпусная база данных коннекторов: построение системы терминов // Информатика и её применения, 2017. Т. 11. Вып. 1. С. 100–108.
12. *Зацман И. М., Кружков М. Г.* Надкорпусная база данных коннекторов: развитие системы терминов проектирования // Системы и средства информатики, 2018. Т. 28. № 4. С. 156–167.
13. *Егорова А. Ю., Зацман И. М., Мамонова О. С.* Надкорпусные базы данных в лингвистических проектах // Системы и средства информатики, 2019. Т. 29. № 3. С. 77–91.
14. *Инькова О. Ю.* Надкорпусная база данных как инструмент формальной вариативности коннекторов // Компьютерная лингвистика и интеллектуальные технологии: по мат-лам ежегодной Международ. конф. «Диалог». — М.: РГГУ, 2018. Вып. 17(24). С. 240–253.

15. Добровольский Д. О., Кретов А. А., Шаров С. А. Корпус параллельных текстов: архитектура и возможности использования // Национальный корпус русского языка: 2003–2005 / Отв. ред. В. А. Плунгян. — М.: Индрик, 2005. С. 263–296.
16. Егорова А. Ю., Зацман И. М., Косарик В. В., Нуриев В. А. Нестабильность нейронного машинного перевода // Системы и средства информатики, 2020. Т. 30. № 2. С. 124–135.
17. Егорова А. Ю., Зацман И. М., Кружков М. Г., Нуриев В. А. Машинный перевод: индикаторная оценка результатов обучения искусственной нейронной сети // Системы и средства информатики, 2020. Т. 30. № 4. С. 124–137.
18. 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.
19. Гончаров А. А., Бунтман Н. В., Нуриев В. А. Ошибки в машинном переводе: проблемы классификации // Системы и средства информатики, 2019. Т. 29. № 3. С. 92–103.

Поступила в редакцию 09.03.21

INDICATOR-BASED EVALUATION OF MACHINE TRANSLATION INSTABILITY

A. Yu. Egorova, I. M. Zatsman, M. G. Kruzchkov, and V. A. Nuriev

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

Abstract: The paper presents data collected by tracking performance of a neural machine translation (NMT) engine and results of translation errors analysis. Indicator-based evaluation of NMT instability was carried out as a part of an experiment that involved 250 Russian text fragments. Each month for the duration of one year, these fragments were translated into French using the Google Translate NMT engine. The translations were recorded and annotated in a supracorpora database; the annotations include types of translation errors found in the translations by language experts. This procedure resulted in a series of 12 annotated translations for each of the 250 Russian fragments. The annotations include not only the types of errors found in the translations but also the types of NMT instability which indicate dynamics of translation quality or lack thereof. The paper aims to provide comparative analysis of Google Translate performance that takes into account the temporal variation aspect.

Keywords: neural machine translation (NMT); instability of machine translation; supracorpora database; indicator-based evaluation; linguistic annotation; NMT instability types

DOI: 10.14357/08696527210213

References

1. Banerjee, S., and A. Lavie. 2005. METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. *43rd Annual Meeting of the Association of Computational Linguistics and the Workshop on Intrinsic and Extrinsic Evaluation Measures for MT and/or Summarization Proceedings*. Ann Arbor, MI: The Association for Computational Linguistics. 65–72.
2. Moorkens, J., S. Castilho, F. Gaspari, and S. Doherty, eds. 2018. *Translation quality assessment*. Machine translation: Technologies and applications ser. Cham: Springer International Publishing. Vol. 1. 299 p.
3. Specia, L., C. Scarton, and G. H. Paetzold. 2018. *Quality estimation for machine translation*. Rafael, CA: Morgan & Claypool Pubs. 162 p.
4. Koehn, Ph. 2020. *Neural machine translation*. New York, NY: Cambridge University Press. 408 p.
5. Mikulić, A. 2020. Ljudska evaluacija sustava za neuralno strojno prevođenje (Završni rad). Available at: <https://urn.nsk.hr/urn:nbn:hr:131:516896> (accessed March 9, 2021).
6. Egorova, A. Yu., I. M. Zatsman, M. G. Kruzhkov, and V. A. Nuriev. 2020. Metodika temporal'noy otsenki nestabil'nosti mashinnogo perevoda [The technique allowing for temporal estimation of the machine translation instability]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(3):67–80.
7. Inkova-Manzotti, O. Yu. 2001. *Konnektory protivopostavleniya vo frantsuzskom i ruskom yazykakh. Sopostavitel'noe issledovanie* [Connectors of opposition in French and Russian: A comparative study]. Moscow: Informelektro. 429 p.
8. Natsional'nyy korpus russkogo yazyka [Russian National Corpus]. Available at: <http://www.ruscorpora.ru/> (accessed March 9, 2021).
9. Zaliznyak, A. A., I. M. Zatsman, O. Yu. Inkova, and M. G. Kruzhkov. 2015. Nadkorpusnye bazy dannykh kak lingvisticheskiy resurs [Supracorpora databases as linguistic resource]. *7th Conference (International) "Corpus Linguistics" Proceedings*. St. Petersburg: SPbSU. 211–218.
10. Durnovo, A. A., I. M. Zatsman, and E. Yu. Loshchilova. 2016. Krosslingvisticheskaya baza dannykh dlya annotirovaniya logiko-semanticheskikh otnosheniy v tekste [Cross-lingual database for annotating logical-semantic relations in the text]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 26(4):124–137.
11. Zaliznyak, A. A., I. M. Zatsman, and O. Yu. Inkova. 2017. Nadkorpurnaya baza dannykh konnektorov: postroenie sistemy terminov [Supracorpora database on connectives: Term system development]. *Informatika i ee Primeneniya — Inform. Appl.* 11(1):100–108.
12. Zatsman, I. M., and M. G. Kruzhkov. 2018. Nadkorpurnaya baza dannykh konnektorov: razvitie sistemy terminov proektirovaniya [Supracorpora database of connectives: Design-oriented evolution of the term system]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(4):156–167.
13. Egorova, A. Yu., I. M. Zatsman, and O. S. Mamonova. 2019. Nadkorpurnye bazy dannykh v lingvisticheskikh proektakh [Supracorpora databases in linguistic projects]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(3):77–91.

14. Inkova, O. Yu. 2018. Nadkorpurnaya baza dannykh kak instrument formal'noy variativnosti konnektorov [Supracorpora database as an instrument of the study of the formal variability of connectives]. *Komp'yuternaya lingvistika i intellektual'nye tekhnologii: po mat-lam Mezhdunar. konf. "Dialog"* [Computer Linguistic and Intellectual Technologies: Conference (International) "Dialog" Proceedings]. Moscow. 17(24):240–253.
15. Dobrovolskiy, 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* [Russian National Corpus: 2003–2005]. Ed. V. A. Plungyan. Moscow: Indrik. 263–296.
16. Egorova, A. Yu., I. M. Zatsman, V. V. Kosarik, and V. A. Nuriev. 2020. Nestabil'nost' neyronnogo mashinnogo perevoda [Instability of neural machine translation]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(2):124–135.
17. Egorova, A. Yu., I. M. Zatsman, M. G. Kruzhkov, and V. A. Nuriev. 2020. Mashinnyy perevod: indikatornaya otsenka rezul'tatov obucheniya iskusstvennoy neyronnoy seti [Machine translation: Indicator-based evaluation of training progress in neural processing]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(4):124–137.
18. Zatsman, I. M., O. Yu. Inkova, M. G. Kruzhkov, 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.
19. Goncharov, A. A., N. V. Buntman, and V. A. Nuriev. 2019. Oshibki v mashinnom perevode: problemy klassifikatsii [Machine translation errors: Problems of classification]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(3):92–103.

Received March 9, 2021

Contributors

Egorova Anna Yu. (b. 1991) — junior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; ann.shurova@gmail.com

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

Nuriev Vitaly A. (b. 1980) — Candidate of Science (PhD) in philology, 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; nurieff.v@gmail.com

УСТОЙЧИВОСТЬ ТЕХНОЛОГИИ ПОДДЕРЖКИ КОНКРЕТНО-ИСТОРИЧЕСКИХ ИССЛЕДОВАНИЙ К ПОПЫТКАМ ИСКАЖЕНИЯ ИСТОРИИ

И. М. Адамович¹, О. И. Волков²

Аннотация: Статья продолжает серию работ, посвященных описанию и анализу распределенной технологии поддержки конкретно-исторических исследований (ПКИИ), основанной на принципах краудсорсинга. Данная статья посвящена моделированию процесса информационных атак на сообщество пользователей технологии (СПТ) с целью фальсификации истории. Выработаны численные показатели, отражающие последствия таких атак. Были проведены проверки, доказывающие уязвимость технологии для информационных атак без применения специальных мер противодействия. Были предложены такие меры, основанные на определенных групповых нормах работы с информацией, распределении ролей и использовании специальных фасилитационных техник и описанные в форме алгоритма поведения членов ядра сообщества. Моделирование показало, что предложенные меры противодействия информационным атакам устойчиво эффективны. Это позволило рекомендовать применение предложенных мер в технологии ПКИИ и сделать вывод об устойчивости обновленной технологии к попыткам фальсификации истории.

Ключевые слова: виртуальное сообщество; модель; технология; искажение истории; конкретно-историческое исследование

DOI: 10.14357/08696527210214

1 Введение

Спор о субъективности и объективности историка идет с момента возникновения исторической науки. Однако в последние годы особо актуальной стала проблема фальсификации истории, где под «фальсификацией», в отличие от субъективного мнения историка, следует иметь в виду сознательное искажение исторических фактов в угоду сиюминутным личным или групповым интересам. Фальсификация истории используется в качестве идеологического оружия внешнеполитическими противниками России [1]. К концу первого десятилетия XXI в. на государственном уровне пришло осознание того, что попытки фальсификации

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

истории, преднамеренные искажения фактов, сознательная и заказная ангажированность трактовок представляют собой не только угрозу академической науке, системе ее критериев и оценок. Это угроза национальной безопасности и государственному суверенитету России. Стратегия национальной безопасности Российской Федерации возводит борьбу с фальсификацией истории в ранг государственной политики [2], при этом исторической науке отводится роль одного из основных инструментов предотвращения распространения ненаучных мифологических концепций исторического прошлого [3]. Но в последнее время сама историческая наука в нашей стране снова становится объектом идейной борьбы, в которой задействованы как внутренние, так и внешние силы. Именно фальсификации истории названы наиболее очевидным и опасным противником исторической науки, поскольку они бывают часто замаскированы под объективное историческое исследование [1]. Действительно, иногда историки отбирают факты, группируют и интерпретируют их таким образом, что в результате возникают исторические мифы. Подчас эти мифы используются в интересах определенных социальных, политических групп, научной или политической элиты, государственной власти [4]. В последнее время уязвимость исследовательских коллективов и сообществ для различных форм внешнего давления только увеличивается в связи с вовлечением в исследовательский процесс не только членов профессионального исторического сообщества, но и самых широких слоев непрофессионалов вследствие неуклонно возрастающего интереса к частной, семейной истории [5].

В [6, 7] описана разработанная в ФИЦ ИУ РАН распределенная технология ПККИ, основанная на принципах краудсорсинга (мобилизации ресурсов широкого круга добровольцев посредством информационных технологий). Технология ПККИ включает в себя онлайн-платформу для коммуникации и совместной деятельности исследователей, и сообщество ее пользователей обладает всеми признаками научно-профессионального сетевого сообщества [8]. В связи с потенциальной уязвимостью СПТ ПККИ для попыток сознательного искажения исторической информации необходим анализ устойчивости технологии ПККИ к фальсификации истории. В [9] описан подход к такому анализу, основанный на моделировании СПТ ПККИ и процессов распространения искаженной информации между ее узлами.

2 Описание модели

Существует множество подходов к моделированию процесса распространения информации в социальных сетях. Особую популярность приобрела SIR (susceptible-infectious-recovered) модель, основанная на сходстве процессов распространения информации в виртуальном сообществе и распространения инфекции в популяции [10]. Недостаток данного подхода — усреднение последствий, в то время как фундаментальным источником риска в таких явлениях выступают экстремумы, а не средние значения [11]. Поэтому для изучения СПТ ПККИ

были выбраны методы имитационного моделирования. За основу была взята модель Боллобаша–Риордана [12], описывающая ориентированный граф, в котором предпочтительное присоединение зависит от входящих и исходящих степеней.

При моделировании социальных сетей, взаимного влияния их членов (агентов), динамики их мнений и т. д. возникает необходимость учета факторов (эффектов), имеющих место в реальных социальных сетях [13]. Поэтому для корректного представления процесса распространения информации в СПТ ПКИИ модель поддерживает следующие характеристики агентов.

1. **Роль:** обычный/атакующий агент. Атакующий агент соответствует злоумышленнику, сознательно и активно распространяющему дезинформацию. Обычные агенты становятся объектами этой атаки. Роль есть постоянная характеристика агента.
2. **Позиция:** нормальная, искаженная, нейтральная. Относится к вопросу, мнение по которому атакующий агент пытается исказить (тема атаки). Нормальная позиция означает, что агент имеет сложившееся корректное, неискаженное мнение по данной теме. Искаженная позиция соответствует искаженному мнению, сложившемуся в результате информационной атаки. Нейтральная позиция означает, что у агента нет сложившегося мнения по теме атаки. Позиция агента может меняться во времени.
3. **Уверенность:** степень уверенности агента в состоятельности своей позиции. Может меняться при воздействии мнений окружения агента. При достижении нулевого уровня уверенности агент меняет свою позицию. Ненулевой уровень уверенности агента с нейтральной позицией соответствует инертности его мышления.
4. **Авторитетность:** коэффициент, с которым воздействие агента оказывает влияние на позицию и уверенность соседа. По сути, является коэффициентом ослабления воздействия.
5. **Активность:** переменный коэффициент, отражающий степень воздействия агента на соседей. Исходно равен нулю, но при появлении воздействия на агента со стороны окружения становится положительным и растет по мере роста уровня воздействия со стороны части окружения, имеющей позицию, отличную от позиции агента.
6. **Воздействие:** вычисляемая характеристика, оказывающая влияние на активность, уверенность и в конечном итоге позицию агента:
 - исходящее воздействие агента пропорционально активности и имеет знак, соответствующий позиции (равно нулю при нейтральной позиции);
 - входящее воздействие агента является суммой исходящих воздействий его окружения с учетом авторитетности воздействующих агентов.

Была проведена серия пробных расчетов модели с исходно нейтральной позицией у всех агентов, кроме атакующего. Такая ситуация не соответствует реальному исходному состоянию СПТ ПКИИ, но позволяет оценить адекватность

модели за счет сравнения результатов распространения искаженной информации в сети с результатами, полученными с помощью моделей «диффузии инноваций», адекватность которых и соответствие процессам распространения информации уже доказана [13]. S-образная форма графика доли узлов с искаженной позицией оказалась соответствующей форме логистической кривой, отражающей процесс диффузии инноваций.

Значительный разброс результатов продемонстрировал преимущество построенной модели перед популярной SIR-моделью за счет возможности оценки не только средних последствий, но и экстремумов, которые выступают фундаментальным источником риска.

Анализ структуры сгенерированного графа показал:

- распределение узлов по числу их связей соответствует степенному закону с показателем степени, лежащим в диапазоне $[2, 3]$ и положительной ассортативностью (характером корреляции узлов, описываемым коэффициентом Пирсона), что типично для социальных сетей в целом и сетей сотрудничества ученых в частности [14];
- до 60% узлов сети могут быть недоступны для информационной атаки, что, как показано в [9], соответствуют реальным научно-профессиональным сетевым сообществам.

Далее под уязвимыми узлами будем понимать узлы графа, потенциально доступные для информационной атаки. Зараженным узлом будем называть узел, сменивший в результате информационной атаки свою позицию на негативную.

3 Результаты моделирования информационной атаки на сообщество пользователей технологии поддержки конкретно-исторических исследований

Была проведена серия экспериментов с моделью с начальными условиями расчета, соответствующими реальному состоянию СПТ ПККИ. Это состояние характеризуется наличием у ряда агентов ранее сформированной нормальной позиции по теме атаки. Каждый эксперимент состоял из серии расчетов с фиксированной начальной долей узлов с нормальной позицией, обозначаемой далее как PositiveProc. По результатам каждого эксперимента предметом анализа были следующие показатели:

- конечная доля зараженных узлов, вычисляемая как среднее отношение числа зараженных узлов к числу уязвимых узлов по состоянию на завершение периода моделирования и обозначаемая далее как EndProc;
- скорость роста заражения, обозначаемая далее как Rate90 и вычисляемая как средний номер шага модельного времени, выраженного в долях от всего периода моделирования, в который доля зараженных узлов достигла 90% от доли зараженных узлов в конце периода моделирования;

Таблица 1 Результаты 1-й серии экспериментов

№ эксперимента	PositiveProc	EndProc	Rate90	V50
1	0%	1,00	0,33	1,00
2	2%	0,80	0,29	0,79
3	5%	0,76	0,34	0,79
4	10%	0,42	0,27	0,46
5	20%	0,12	0,15	0,12
6	50%	0,01	0,06	0,00

- уязвимость технологии ПКИИ, понимаемая как вероятность того, что доля зараженных узлов превысит порог в 50% на каком-либо шаге расчета, и обозначаемая далее как V50.

Результаты 1-й серии экспериментов приведены в табл. 1. Видно, что по мере роста доли узлов (агентов) с заранее сформированной позицией по теме атаки (PositiveProc) негативные последствия информационных атак (EndProc и V50) снижаются. Но вероятность значительных последствий сохраняется вплоть до достижения PositiveProc уровня 50%, что в реальных условиях трудно ожидать. Это означает, что технология ПКИИ уязвима для информационных атак и требуются специальные меры противодействия им. Также следует отметить, что по мере роста PositiveProc показатель Rate90 также снижается, но характер снижения нелинеен, из чего следует, что на скорость роста заражения оказывает значительное влияние топология социальной сети.

4 Описание подхода к противодействию информационным атакам

Как было показано выше, технология ПКИИ потенциально уязвима для информационных атак, имеющих целью фальсификацию истории. Следовательно, в технологии должны быть предусмотрены меры по противодействию таким информационным атакам. Как показано в [15], противодействие распространению ложной информации в сетевых сообществах может опираться как на машинные алгоритмы, так и на более сложные социальные технологии, повышающие рефлексивность сообществ (способность сообществ анализировать эффективность своей деятельности).

Первый подход уже реализуется, но его применение затруднительно в условиях технологии ПКИИ, поскольку фальсификации истории часто замаскированы под объективное историческое исследование и их выявление по формальным признакам невозможно. Второй подход, основанный на социальных технологиях поддержания коллективной рефлексивности, критического мышления и коллективного эмоционального интеллекта, еще только предстоит разработать. Как показано в [15], он должен быть направлен на повышение устойчивости сете-

вых сообществ к «информационной войне» за счет формирования определенных групповых норм работы с информацией, распределения ролей и использования специальных фасилитационных техник (техники управления, отличающиеся от простого управления тем, что их способ недирективный, т. е. не выходит за рамки самоорганизации управляемой системы).

Меры противодействия информационным атакам на СПТ ПККИ должны строиться на специфике данного сообщества как научно-профессионального. В [16] приведен анализ профессиональных сообществ и определены значения ряда их измеряемых характеристик. Из анализа следует, что для профессиональных сообществ характерно наличие развитого ядра (наиболее активной части сообщества), ориентированного не только на личный профессиональный рост и получение результатов в интересах своих исследований, но и на организационные усилия и обмен знаниями в интересах всего сообщества. К ядру сообщества относят себя больше 66% респондентов, попавших в выборку. При этом около половины опрошенных участников занимаются организаторской деятельностью в своих сообществах. Моральное удовлетворение от трансляции своего опыта часто испытывают 77,8% респондентов. В качестве основной причины, по которой люди принимают участие в деятельности профессиональных сообществ, около 5% респондентов назвали «чтобы делиться опытом». Из этого следует, что не менее 5% участников сообщества готовы добровольно взять на себя дополнительную роль в сообществе, связанную с противодействием информационным атакам. Представляется целесообразным связать эту роль со следующим алгоритмом действий.

1. На этапе формирования сообщества:

- создается карта ядра сообщества с указанием контакта каждого члена и областей его специализации;
- при появлении нового члена, добровольно принимающего роль участника противодействия информационным атакам, его данные добавляются в карту ядра.

2. При обнаружении факта информационной атаки:

- при обнаружении любым членом ядра сообщества информационных сообщений, предположительно содержащих фальсифицированную информацию и являющихся частью информационной атаки, сообщение пересылается тем членам ядра, специализация которых позволяет профессионально оценить данное сообщение (далее — эксперты);
- в случае подтверждения предположений об информационной атаке эксперты совместно формируют сообщение или серию сообщений с обоснованной критикой фальсификации (далее — опровержение);
- все члены ядра обеспечивают распространение опровержения в сообществе посредством его репоста.

Алгоритм предполагает быстрое обнаружение атаки и одновременное массовое противодействие ей. Эффективность данного алгоритма может быть проверена экспериментально с помощью описанной выше модели.

5 Результаты моделирования процессов противодействия атакам

Была проведена вторая серия экспериментов с моделью, целью которых было оценить влияние активного ядра, действующего в соответствии с вышеописанным алгоритмом противодействия информационной атаке, с различным объемом ядра, выраженного в процентах от всего объема сообщества и далее обозначаемого как CoreProc. Для выявления чистого влияния параметра CoreProc на эффективность противодействия параметр PositiveProc был установлен равным нулю. Оценивались те же показатели, что и в первой серии экспериментов. Результаты проверок приведены в табл. 2.

Таблица 2 Результаты 2-й и 3-й серии экспериментов

№ серии	№ эксперимента	Количество атакующих	CoreProc	PositiveProc	EndProc	Rate90	V50
2	7	1	0,025%	0%	0,94	0,33	0,94
	8		0,050%		0,09	0,13	0,01
3	9	25	0,025%	5%	0,77	0,25	0,80
	10		0,050%		0,09	0,17	0,11
	11		0,100%		0,06	0,08	0,00

Из результатов экспериментов второй серии видно, что даже при крайне незначительном объеме ядра (0,05% от всего сообщества) СПТ ПККИИ способно успешно противостоять информационной атаке. Это говорит о высокой эффективности предложенного алгоритма. Но также требует проверки степень снижения эффективности противодействия по мере роста интенсивности информационной атаки. Для этого была проведена третья серия экспериментов с различными значениями CoreProc, но с 25-кратным увеличением интенсивности атаки (25 одновременно атакующих узлов). Параметр PositiveProc для реалистичности был установлен равным 5%.

Из результатов экспериментов третьей серии видно, что, несмотря на значительное увеличение интенсивности атаки, снижение эффективности противостояния было невелико и при объеме ядра 0,1% от всего СПТ ПККИИ оно способно успешно противостоять информационной атаке даже в этих условиях.

6 Выводы

Результаты моделирования показали, что без применения специальных мер противодействия технология ПККИИ потенциально уязвима для информационных атак, имеющих целью фальсификацию истории.

Моделирование также показало, что предложенные меры противодействия информационным атакам, основанные на определенных групповых нормах работы с информацией, распределении ролей и использовании специальных фасилитационных техник и описанные в форме алгоритма поведения членов ядра СПТ ПККИ, устойчиво эффективны. Это позволяет рекомендовать применение предложенных мер в технологии ПККИ и сделать вывод об устойчивости обновленной технологии к попыткам фальсификации истории.

Литература

1. *Алексеев С. В., Плотникова О. А.* Мифы и фальсификации в российской истории // *Знание. Понимание. Умение*, 2015. № 1. С. 162–171.
2. *Якеменко Б. Г.* Фальсификации российской истории — нужно ли с ними бороться, для чего и какими методами? // *Вестник Российского университета дружбы народов*. Сер.: История России, 2012. № 4. С. 131–135.
3. *Вяземский Е. Е.* Проблема фальсификации истории России и общее историческое образование: теоретические и практические аспекты // *Проблемы современного образования*, 2012. № 1. С. 28–43.
4. *Криворученко В. К.* Борьба с фальсификацией истории на постсоветском пространстве: история или политика? // *Вестник Российского университета дружбы народов*. Сер.: История России, 2012. № 4. С. 121–130.
5. *Помникова А. Ю.* Семейная история в дискурсивном пространстве // *Вестник Мининского университета*, 2019. Т. 7. № 1(26). Ст. 9. 22 с.
6. *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.
7. *Адамович И. М., Волков О. И.* Единая технология поддержки конкретно-исторических исследований // *Системы и средства информатики*, 2019. Т. 29. № 1. С. 194–205. doi: 10.14357/08696527190116.
8. *Каменский Е. Г., Гримов О. А.* Сетевые сообщества в социальных сетях как фактор развития личностной субъектности // *Вестник Нижегородского университета им. Н. И. Лобачевского*. Сер.: Социальные науки, 2014. № 2(34). С. 62–67.
9. *Адамович И. М., Волков О. И.* Модель сообщества пользователей технологии поддержки конкретно-исторических исследований // *Системы и средства информатики*, 2021. Т. 31. № 1. С. 145–156.
10. *Васенин В. А., Афонин С. А., Панюшкин Д. С.* Модели распространения информации в социальных сетях // *Программная инженерия*, 2014. № 2. С. 33–42.
11. *Cirillo P., Taleb N.* Tail risk of contagious diseases // *Nat. Phys.*, 2020. Vol. 16. P. 606–613.
12. *Bollobás B., Borgs C., Chayes T., Riordan O. M.* Directed scale-free graphs // *14th Annual ACM-SIAM Symposium on Discrete Algorithms Proceedings*. — New York, NY, USA: ACM, 2003. P. 132–139.
13. *Губанов Д. А., Новиков Д. А., Чхартушвили А. Г.* Модели влияния в социальных сетях // *Управление большими системами*, 2009. № 27. С. 205–281.

14. *Евин И. А., Хабибуллин Т. Ф.* Социальные сети // Компьютерные исследования и моделирование, 2012. Т. 4. № 2. С. 423–430.
15. *Мухеев Е. А., Нестик Т. А.* Дезинформация в социальных сетях: состояние и перспективы психологических исследований // Социальная психология и общество, 2018. Т. 9. № 2. С. 5–20.
16. *Иванов О. В., Карлюкова О. С.* Профессиональные сообщества в России: количественный анализ // e-LearningWorld, 2007. № 1(17). С. 90–93.

Поступила в редакцию 11.11.20

RESISTANCE OF TECHNOLOGY OF CONCRETE HISTORICAL INVESTIGATION SUPPORT TO ATTEMPTS OF HISTORY DISTORTION

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 119133, Russian Federation

Abstract: The article continues a series of works devoted to the description and analysis of distributed technology of concrete historical investigation support based on the principles of crowdsourcing. This article is devoted to modeling the process of information attacks on the community of technology users in order to falsify the history. Some numerical indicators reflecting the consequences of such attacks have been developed. The tests proved the vulnerability of the technology to information attacks in the absence of special countermeasures. Such measures based on certain group norms of information processing, the distribution of roles, and the use of special facilitation techniques and described in the form of community members activity algorithm were proposed. Modeling has shown that the proposed measures of resistance to information attacks are stably effective. It allowed us to recommend the usage of the proposed measures in the technology of concrete historical investigation support and to draw a conclusion about the sustainability of the updated technology to attempts to falsify the history.

Keywords: virtual community; model; technology; distortion of history; concrete historical investigation

DOI: 10.14357/08696527210214

References

1. Alekseev, S. V., and O. A. Plotnikova. 2015. Mify i fal'sifikatsii v rossiyskoy istorii [Myths and falsifications in the history of Russia]. *Znanie. Ponimanie. Umenie* [Knowledge. Understanding. Skill] 1:162–171.

2. Yakemenko, B. G. 2012. Fal'sifikatsii rossiyskoy istorii — nuzhno li s nimi borot'sya, dlya chego i kakimi metodami? [Falsifications of Russian history — should we fight against them, for which purpose and by which means?]. *RUDN J. Russian History* 4:131–135.
3. Vyazemsky, Ye. Ye. 2012. Problema fal'sifikatsii istorii Rossii i obshchee istoricheskoe obrazovanie: teoreticheskie i prakticheskie aspekty [Problem of falsification of Russian history and the general historical education: Theoretical and practical aspects]. *Problemy sovremennogo obrazovaniya* [Problems of Modern Education] 1:28–43.
4. Krivoruchenko, V. K. 2012. Bor'ba s fal'sifikatsiyey istorii na postsovetskom prostranstve: istoriya ili politika? [Counteraction of falsification of history in former USSR: History or politics?]. *RUDN J. Russian History* 4:121–130.
5. Pomnikova, A. Yu. 2019. Semeynaya istoriya v diskursivnom prostranstve [Family stories in different types of discourse]. *Vestnik of Minin University* 7(1):9. 22 p.
6. Adamovich, I. M., and O. I. Volkov. 2016. Tekhnologiya raspredelenogo avtomatizirovannogo analiza istoricheskikh tekstov [The distributed automated technology of historical texts analysis]. *Sistemy i Sredstva Informatiki - Systems and Means of Informatics* 26(3):148–161. doi: 10.14357/08696527160311.
7. Adamovich, I. M., and O. I. Volkov. 2019. Edinaya tekhnologiya podderzhki konkretno-istoricheskikh issledovaniy [Unified technology of concrete historical research support]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(1):194–205. doi: 10.14357/08696527190116.
8. Kamensky, E. G., and O. A. Grimov. 2014. Setevyye soobshchestva v sotsial'nykh setyakh kak faktor razvitiya lichnostnoy sub'yektnosti [Network communities in social networks as a factor in the development of personal subjectivity]. *Vestnik of Lobachevsky State University of Nizhni Novgorod. Sociol sciences ser.* 2(34):62–67.
9. Adamovich, I. M., and O. I. Volkov. 2021. Model' soobshchestva pol'zovateley tekhnologii podderzhki konkretno-istoricheskikh issledovaniy [The model of community of concrete historical investigation support technology users]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 31(1):145–156.
10. Vasenin, V. A., S. A. Afonin, and D. S. Panushkin. 2014. Modeli rasprostraneniya informatsii v sotsial'nykh setyakh [Models of information dissemination in social networks]. *Programmnaya inzheneriya* [Software Engineering] 2:33–42.
11. Cirillo, P., and N. Taleb. 2020. Tail risk of contagious diseases. *Nat. Phys.* 16:606–613.
12. Bollobás, B., C. Borgs, T. Chayes, and O. M. Riordan. 2003. Directed scale-free graphs. *14th Annual ACM-SIAM Symposium on Discrete Algorithms Proceedings*. New York, NY: ACM. 132–139.
13. Gubanov, D. A., D. A. Novikov, and A. G. Chkhartishvili. 2009. Modeli vliyaniya v sotsial'nykh setyakh [Models of influence in social networks]. *Large-Scale Systems Control* 27:205–281.
14. Yevin, I. A., and T. F. Khabibullin. 2012. Sotsial'nye seti [Social networks]. *Komp'yuternye issledovaniya i modelirovanie* [Computer Research and Modeling] 4(2):423–430.
15. Mikheev, E. A., and T. Nestik. 2018. Dezinformatsiya v sotsial'nykh setyakh: sostoyanie i perspektivy psikhologicheskikh issledovaniy [Disinformation in social networks: state and perspectives of psychological research].

Current state and perspective research directions]. *Sotsial'naya psikhologiya i obshchestvo* [Social Psychology and Society] 9(2):5–20.

16. Ivanov, O. V., and O. S. Karlyukova. 2007. Professional'nye soobshchestva v Rossii: kolichestvennyy analiz [Professional communities in Russia: Quantitative analysis]. *e-LearningWorld* 1(17):90–93.

Received November 11, 2020

Contributors

Adamovich Igor M. (b. 1934)— Candidate of Science (PhD) in technology, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; Adam@amsd.com

Volkov Oleg I. (b. 1964) — leading programmer, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; Volkov@amsd.com

СИТУАЦИОННАЯ ЦИФРОВИЗАЦИЯ ТОВАРНО-ДЕНЕЖНОГО ОБРАЩЕНИЯ

А. В. Ильин¹, В. Д. Ильин²

Аннотация: Предложена модель ситуационной цифровизации системы товарно-денежного обращения (*тдо*-системы), в которой спецификации товаров, платежей и других составляющих системы представлены зашифрованными цифровыми кодами. Управление товарно-денежным обращением осуществляется посредством систем обязательных и ориентирующих требований, зафиксированных в ситуационных цифровых договорах, определяющих отношения заказчиков и поставщиков. Процесс выполнения каждого договора представлен определенными в нем ситуациями. Данные, определяющие портреты ситуаций, поставляются цифровыми двойниками, обслуживающими договор. Эти данные передаются на вход программы договора, выполняемой на компьютерных устройствах договорившихся сторон. Перемещение товаров и платежей сопровождается документирующим перемещением соответствующих кодов. Каждое перемещение отражается в памяти цифровой системы учета величин имущественного права, представленной совокупностью унифицированных счетов поставщиков и заказчиков. Мультивалютные счета открываются и программно контролируются с помощью онлайн-сервисов банка-регулятора. Для каждого товара сделка может быть совершена в любой валюте из списка, являющегося пересечением множеств валют, разрешенных банками-регуляторами государств, к которым относятся участники сделки.

Ключевые слова: система товарно-денежного обращения; ситуационный цифровой договор; система учета величин имущественного права; унифицированный мультивалютный счет; онлайн-сервис

DOI: 10.14357/08696527210215

1 Введение

Цифровизация *системы товарно-денежного обращения* входит в число ключевых задач цифровизации хозяйственной деятельности [1–3]. Наряду с унификацией цифровой маркировки товаров (для документируемого мониторинга их перемещения) и решением других актуальных задач цифровизации *тдо*-системы необходимо преобразовать онлайн-сервисы банковской системы. Их целесообразно осуществлять с помощью *персональных электронных банков (пэбов)* и *корпоративных электронных банков (кэбов)* [4–6], работающих под

¹Государственный научно-исследовательской 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

управлением Центрального банка (глава Сбербанка Герман Греф считает, что «со временем банковская система станет одноуровневой» (счета клиентов будут открыты непосредственно в Центральном банке)¹).

Одно из опасных следствий неадекватности нынешней банковской системы. На современном этапе цифровизации экономики серьезную опасность для государственного *тдо*-регулирования представляет незаконное использование биткойна и/или других криптовалют как средств платежа и/или инвестирования. Неподконтрольную национальным банкам-регуляторам эмиссию криптовалют осуществляют анонимные пользователи криптовалютных блокчейн-сетей, используя для этого собственные компьютерные устройства (и получая необлагаемую налогом прибыль). Перемещение криптовалютных сумм между кошельками пользователей криптовалютных сетей также неподконтрольно банкам-регуляторам.

В связи с этим в России с 1 января 2021 г. действует закон о цифровых финансовых активах и цифровой валюте². «Цифровой валютой признается совокупность электронных данных (цифрового кода или обозначения), содержащихся в информационной системе, которые предлагаются и (или) могут быть приняты в качестве средства платежа, не являющегося денежной единицей Российской Федерации, денежной единицей иностранного государства и (или) международной денежной или расчетной единицей, и (или) в качестве инвестиций и в отношении которых отсутствует лицо, обязанное перед каждым обладателем таких электронных данных, за исключением оператора и (или) узлов информационной системы, обязанных только обеспечивать соответствие порядка выпуска этих электронных данных и осуществления в их отношении действий по внесению (изменению) записей в такую информационную систему по ее правилам».

В переводе с «юридического»: использование криптовалют (как средств платежа или инвестирования) запрещено из-за отсутствия «лиц, обязанных перед обладателями» криптовалютного достояния. При использовании криптовалют в качестве «цифровых финансовых активов» оператор «информационной системы», в которой осуществляется их выпуск, обязан обеспечить хранение данных о сделках с цифровыми финансовыми активами, а также об участниках таких сделок (не менее пяти лет с даты совершения соответствующих сделок)».

Поскольку это требование противоречит основополагающим принципам неподконтрольности и анонимности участников криптовалютных сетей, использование нынешних криптовалют в качестве «цифровых финансовых активов» в России также противозаконно.

Инфраструктурная платформа. На современном этапе инфраструктурная платформа цифровизации *тдо*-системы включает технологии *цифровых двойни-*

¹Греф предрек похороны современной банковской системы. <https://ria.ru/20160521/1437470618.html>.

²Федеральный закон от 31.07.2020 г. № 259-ФЗ «О цифровых финансовых активах, цифровой валюте и о внесении изменений в отдельные законодательные акты Российской Федерации». <http://www.kremlin.ru/acts/bank/45766>.

ков (Digital Twins¹), M2M-технологии (M2M — Machine-to-Machine) [7, 8], технологии облачных вычислений (Cloud Computing) и электронных сервисов [9–12], интернета вещей (Internet of Things, IoT) [13].

Запись формул и выделение фрагментов текста. Для выделения определений, замечаний и примеров используются средства языка TSM-комплекса (TSM — textual symbolic modeling), разработанного для формализованного описания текстовых моделей².

В статье применены следующие средства выделения фрагментов текста:

□ ⟨фрагмент описания⟩ □ ≈ утверждение (определение, аксиома и др.) (здесь и далее символ ≈ заменяет слово «означает»);

◇ ⟨фрагмент описания⟩ ◇ ≈ замечание;

○ ⟨фрагмент описания⟩ ○ ≈ пример.

Курсивом выделены первые вхождения названий понятий и фрагменты описания, к которым автор хочет привлечь внимание.

Обсуждаемые результаты. В статье, которая служит обобщающим и уточняющим продолжением [6], представлены результаты, полученные при выполнении научно-исследовательской работы «Моделирование социальных, экономических и экологических процессов» (№ 0063-2016-0005), выполняемой в соответствии с государственным заданием ФАНО России для Федерального исследовательского центра «Информатика и управление» РАН. Концептуальные основы предложенной модели, обсуждаемые в статье, рассматриваются как часть методологического обеспечения проекта ситуационной цифровизации *тдо*-системы.

2 Цифровая маркировка товаров

Каждый продукт (произведенный или ввезенный в страну), получив цифровую маркировку, становится *товаром*, перемещение которого контролируется в процессе *товарно-денежного обращения* (*тдо*-товаром). *Цифровая маркировка* является закодированной спецификацией *тдо*-товара, содержащей его наименование, дату выпуска и другие сведения.

3 Ситуационный цифровой договор

Все *тдо*-товары разделены на два класса [14]: поставляемых в рамках сделок, совершающихся по договорам, и поставляемых на бездоговорной основе (○ товары, приобретаемые покупателями в магазинах розничных торговых сетей ○).

¹The New Age of Manufacturing: Digital Twin Technology & IIoT. <https://medium.com/@lior.kitain/the-new-age-of-manufacturing-digital-twin-technology-iiot-494acee5572a>.

²Ильин В. Д. Символьное моделирование // Большая российская энциклопедия (электронная версия). http://dev.bigenc.ru/technology_and_technique/text/4010980.

Рассматриваемая далее *ситуационная модель цифрового договора* основана на идее *смайт-контракта*, предложенной Ником Сабо [15], и авторской методологии *ситуационной информатизации* [16, 17].

□ *Ситуационный цифровой договор (сцд) выполняется с помощью* специальной программно-аппаратно реализуемой интеллектуальной системы (*сцд-системы*), которая содержит описание *тдо-сделки* и алгоритмически определенные *обязательные* и *ориентирующие* требования к участникам сделки. В процессе контроля исполнения договора *сцд-система* обменивается данными с созданными ею обучаемыми цифровыми двойниками, отслеживающими события (представленные *портретами ситуаций* [16, 17]), определенные в договоре. *Обязательное требование* имеет вид $s^m \rightarrow a^m$, где s^m — спецификация события, при наступлении которого определенный в договоре участник (или участники) сделки должен выполнить действие a^m . *Ориентирующее требование* — выражение вида $s^{or} \rightarrow i^{or}$, где s^{or} — спецификация события, при наступлении которого определенный участник (или участники) должен выполнить одно из действий в соответствии с инструкцией i^{or} . □

4 Цифровая система учета величин имущественного права

□ *Цифровая система учета величин имущественного права (ип-система)* рассматривается как реализуемая на основе информационно-безопасных банковских онлайн-сервисов система документального представления денежной и неденежной составляющих имущественных статусов участников *тдо-сделок*. Денежная составляющая выражена значениями сумм *нормализованных денег* в разделах *уникальных мультивалютных счетов* участников *тдо-сделок*, а неденежная — документами, подтверждающими право собственности на недвижимость, транспорт и другое имущество (которое при необходимости может рассматриваться как залоговое). □

◇ Уникальные мультивалютные счета размещены в *пэбах* (для физических лиц) и в *кэбах* (для юридических лиц) [6]. ◇

□ *Нормализованные деньги (нд)* — универсальное электронное средство количественного документирования имущественных отношений, удостоверяемых государством. Предназначено для представления стоимости товаров и имущественных статусов участников экономической деятельности, для оплаты товаров, инвестирования и накопления богатства. *Универсальность нд* выражается в применимости во всех внутренних и внешних экономических сделках, разрешенных законом государства, под юрисдикцией которого функционирует экономическая система [4–6]. □

5 Базовые технологии цифровизации товарно-денежного обращения

В состав базовых технологий цифровизации товарно-денежного обращения входят *технологии цифровой маркировки товаров, назначенных платежей, пла-*

тежного и товарного кредитования, онлайн-банкинга с использованием кэбов и пэбов, принадлежащих участникам тдо-сделок [3, 4].

□ *Технология назначенных платежей* — совокупность методов, средств и правил гарантированной оплаты заказанных товаров, требования к реализации которой определены в *сид* о тдо-сделке. Сумма назначенного платежа переводится на счет получателя сразу после того, как заказчик подтвердил исполнение заказа. В случае неисполнения заказа назначенный платеж отменяется [18]. □

В концепции поливалютного рынка *нормализованного экономического механизма* [4, 5] *технологии внешнеэкономических тдо-сделок* основаны на предположении о том, что для любой пары стран на каждом отрезке времени определены некоторые множества товаров, множество применимых для расчетов валют, таможенные правила и правила внешнеторговых сборов. Для каждого типа товара участники тдо-сделки имеют возможность совершать назначенные платежи в выбранной ими валюте (из списка, являющегося пересечением множеств валют, разрешенных банками-регуляторами государств, к которым относятся участники тдо-сделки (○ в российских рублях и юанях ○)).

Одной из наиболее актуальных технологий товарного кредитования становится □ *технология электронной долговой торговли* [19], при которой отсроченная часть оплаты товара оформляется как *долг покупателя продавцу*. □ ◇ Применительно к долгам, образовавшимся в результате продажи приоритетных товаров, могут действовать правила досрочного возврата долгов продавцам из средств долгового отдела банка-регулятора. Таким способом банк-регулятор может участвовать в государственных программах поддержки производства и продажи приоритетных товаров. ◇

Важной составляющей технологии электронной долговой торговли является □ *долговая денежная эмиссия*, выполняемая банком-регулятором только тогда, когда сумма возвращенных покупателями долгов меньше очередной долговой суммы, запрошенной для возврата продавцу. Эмитируемая сумма равна разности запрошенной долговой суммы и суммы на счету долгового отдела банка-регулятора. □

◇ Правила, реализуемые в технологии, исключают возможность эмиссии денежных сумм, не обеспеченных товарами. ◇

◇ На поливалютном рынке тдо-сделок технология электронной долговой торговли служит средством, стимулирующим продажи приоритетных товаров с оплатой в национальной валюте, так как возможность в относительно короткие сроки получить долговую часть стоимости проданного товара существует, если он был оплачен национальной валютой. ◇

◇ Целесообразны отношения координации между странами (при выработке и реализации схем контроля исполнения указанных правил) и нецелесообразны любые глобальные регуляторы, ограничивающие свободу экономического выбора участников тдо-сделок. ◇

Технологии платежного кредитования включают технологии краткосрочного и долгосрочного кредитования поставщиков товаров. Краткосрочное

платежное кредитование может быть представлено частями полной оплаты заказанного товара, поступающими на счет поставщика до исполнения заказа (○ при поставке товара, представляющего собой изделие, изготавливаемое по индивидуальному заказу ○). *Долгосрочное платежное кредитование* может осуществляться в обмен на ценовые скидки или иные преимущества (○ выражающиеся в сокращенных сроках поставок ○).

6 Заключение

Определены основные понятия и базовые технологии *ситуационной цифровизации тдо-системы*.

Каждая *тдо-сделка* совершается в соответствии с *сид*, выполнение которого осуществляется с помощью специальной интеллектуальной системы (*сид-системы*), содержащей описание сделки и *обязательные* и *ориентирующие* требования к действиям участников.

Выполнение договора представлено спецификациями событий, для каждого из которых определен *портрет ситуации* [16, 17], хранящийся в памяти *сид-системы* (как образец для сравнения).

Портреты ситуаций, возникающих при выполнении договора, строятся *сид-системой* на основе данных, поставляемых *цифровыми двойниками*. Для идентификации событий, определенных в договоре, построенные портреты ситуаций сравниваются с образцами, хранящимися в памяти *сид-системы*.

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

Уникальные мультивалютные счета участников *тдо-сделок*, открываемые и программно контролируемые с помощью онлайн-сервисов банка-регулятора, размещаются в *нэбах* (для физических лиц) и в *кэбах* (для юридических лиц) [4–6].

К арсеналу базовых технологий цифровизации товарно-денежного обращения отнесены *технологии цифровой маркировки товаров, онлайн-банкинга, назначенных платежей* [18], *товарного и платежного кредитования*. Особая роль в выполнении государственных программ поддержки производства и поставок приоритетных товаров принадлежит *технологии электронной долговой торговли* [19].

Литература

1. The New Digital Economy: How it will transform business. — Oxford Economics, 2015. 34 p. <http://www.pwc.com/mt/en/publications/assets/the-new-digital-economy.pdf>.
2. G20 digital economy development and cooperation initiative // G20 Summit, 2016. <http://en.kremlin.ru/supplement/5111>.

3. Цифровая экономика Российской Федерации: Программа, утвержденная распоряжением Правительства Российской Федерации от 28 июля 2017 г. № 1632-р.
4. *Ilyin A. V., Ilyin V. D.* Towards a normalized economic mechanism based on E-services // *Agris on-line Papers in Economics and Informatics*, 2014. Vol. 6. Iss. 3. P. 39–49.
5. *Ilyin A. V., Ilyin V. D.* The normalized economic mechanism in the digital environment // *Int. J. Open Information Technologies*, 2019. Vol. 7. Iss. 12. P. 77–83.
6. *Ильин А. В., Ильин В. Д.* Персональные и корпоративные электронные банки в технологиях товарно-денежного обращения // *Системы и средства информатики*, 2020. Т. 30. № 2. С. 187–194.
7. *Kim R. Y.* Efficient wireless communications schemes for machine to machine communications // *Comm. Com. Inf. Sc.*, 2011. Vol. 181. No. 3. P. 313–323.
8. *Lien S. Y., Liao T. H., Kao C. Y., et al.* Cooperative access class barring for machine-to-machine communications // *IEEE T. Wirel. Commun.*, 2012. Vol. 11. No. 1. P. 27–32.
9. *Armburst M., Fox A., Griffith R., et al.* A view of cloud computing // *Commun. ACM*, 2010. Vol. 53. No. 4. P. 50–58.
10. *Wei Y., Blake M. B.* Service-oriented computing and cloud computing: Challenges and opportunities // *IEEE Internet Comput.*, 2010. Vol. 14. No. 6. P. 72–75.
11. *Rogers O., Cliff D.* A financial brokerage model for cloud computing // *J. Cloud Computing*, 2012. Vol. 1. No. 1. P. 1–12.
12. *Jede A., Teuteberg F.* Understanding socio-technical impacts arising from software as-a-service usage in companies // *Bus. Inf. Syst. Eng.*, 2016. Vol. 58. No. 3. P. 161–176.
13. *Perera C., Liu C. H., Jayawardena S.* The emerging Internet of Things marketplace from an industrial perspective: A survey // *IEEE T. Emerging Topics Computing*, 2015. Vol. 3. No. 4. P. 585–598.
14. *Ильин А. В., Ильин В. Д.* Ситуационная модель универсального договора в цифровой среде // *Системы и средства информатики*, 2020. Т. 30. № 3. С. 154–162.
15. *Szabo N.* Formalizing and securing relationships on public networks // *First Monday*, 1997. <https://firstmonday.org/article/view/548/469>.
16. *Ильин В. Д.* Основания ситуационной информатизации. — М.: Наука. Физматлит, 1996. 180 с.
17. *Ильин В. Д., Гавриленко Ю. В., Ильин А. В., Макаров Е. М.* Математические средства ситуационной информатизации. — М.: Наука. Физматлит, 1996. 88 с.
18. *Ильин В. Д.* Технология назначенных платежей в среде цифровых двойников // *Системы и средства информатики*, 2018. Т. 28. № 3. С. 227–235.
19. *Ilyin A. V., Ilyin V. D.* E-trade with direct lending and normalized money // *Agris on-line Papers in Economics and Informatics*, 2015. Vol. 7. Iss. 4. P. 57–64.

Поступила в редакцию 11.03.21

SITUATIONAL DIGITALIZATION OF COMMODITY-MONEY CIRCULATION

A. V. Ilyin¹ and V. D. Ilyin²

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

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

Abstract: A model of situational digitalization of the commodity-money circulation system is proposed, where the specifications of goods, payments, and other components of the system are represented by encrypted digital codes. The management of commodity-money circulation is carried out through the systems of mandatory and orienting requirements fixed in situational digital contracts that determine the relationships between customers and suppliers. The implementation of each contract is represented by the situations defined there. Data defining the portraits of situations are supplied by digital twins serving the contract. These data are going to the input of the contract program executed on the computer devices of the contracting parties. The movement of goods and payments is accompanied by the documentary supported movement of the corresponding codes. Each operation is reflected in the memory of the digital system of recording the values of property rights, represented by a set of unified accounts of suppliers and customers. Multicurrency accounts are opened and programmatically controlled by the online services of the regulating bank. For each commodity, the transaction can be made in any currency from the list which is the intersection of the sets of currencies allowed by the regulating banks of the states to which the transaction parties belong.

Keywords: system of commodity-money circulation; situational digital contract; system for recording the values of property rights; unified multicurrency account; online service

DOI: 10.14357/08696527210215

References

1. Oxford Economics. 2015. The new digital economy: How it will transform business. Available at: <http://www.pwc.com/mt/en/publications/assets/the-new-digital-economy.pdf> (accessed March 24, 2021).
2. G20 digital economy development and cooperation initiative. 2016. G20 Summit. Available at: <http://en.kremlin.ru/supplement/5111> (accessed March 24, 2021).
3. Tsifrovaya ekonomika Rossiyskoy Federatsii: Programma, utverzhennaya rasporyazheniem Pravitel'stva Rossiyskoy Federatsii [Digital economy of the Russian Federation: Program approved by Order No. 1632-r dated July 28, 2017 of the Government of

- the Russian Federation]. Available at: <http://d-russia.ru/wpcontent/uploads/2017/07/programma-tsifrov-econ.pdf> (accessed March 24, 2021).
4. Ilyin, A. V., and V. D. Ilyin. 2014. Towards a normalized economic mechanism based on E-services. *Agris on-line Papers in Economics and Informatics* 6(3):39–49
 5. Ilyin, A. V., and V. D. Ilyin. 2019. The normalized economic mechanism in the digital environment. *Int. J. Open Information Technologies* 7(12):77–83.
 6. Ilyin, A. V., and V. D. Ilyin. 2020. Personal'nye i korporativnye elektronnye banki v tekhnologiyakh tovarno-denezhnogo obrashcheniya [Personal and corporate E-banks in commodity-money circulation technologies]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(2):187–194.
 7. Kim, R. Y. 2011. Efficient wireless communications schemes for machine to machine communications. *Comm. Com. Inf. Sc.* 181(3):313–323.
 8. Lien, S. Y., T. H. Liaw, C. Y. Kao, et al. 2012. Cooperative access class barring for machine-to-machine communications. *IEEE T. Wirel. Commun.* 11(1):27–32.
 9. Armbrust, M., A. Fox, R. Griffith, et al. 2010. A view of cloud computing. *Commun. ACM* 53(4):50–58.
 10. Wei, Y., and M. B. Blake. 2010. Service-oriented computing and cloud computing: Challenges and opportunities. *IEEE Internet Comput.* 14:72–75.
 11. Rogers, O., and D. Cliff. 2012. A financial brokerage model for cloud computing. *J. Cloud Comput.* 1(1):1–12.
 12. Jede, A., and F. Teuteberg. 2016. Understanding socio-technical impacts arising from software as-a-service usage in companies. *Bus. Inf. Syst. Eng.* 58(3):161–176.
 13. Perera, C., C. H. Liu, and S. Jayawardena. 2015. The emerging Internet of Things marketplace from an industrial perspective: A survey. *IEEE T. Emerging Topics Computing* 3(4):585–598.
 14. Ilyin, A. V., and V. D. Ilyin. 2020. Situatsionnaya model' universal'nogo dogovora v tsifrovoy srede [Situational model of universal agreement in a digital environment]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(3):154–162.
 15. Szabo, N. Formalizing and securing relationships on public networks. *First Monday*. Available at: <https://firstmonday.org/article/view/548/469> (accessed March 10, 2021).
 16. Ilyin, V. D. 1996. *Osnovaniya situatsionnoy informatizatsii* [Fundamentals of situational informatization]. Moscow: Nauka, Fizmatlit. 180 p.
 17. 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.
 18. Ilyin, V. D. 2018. Tekhnologiya naznachennykh platezhey v srede tsifrovyykh dvoynikov [Designated payments technology in digital twins environment]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(3):227–235.
 19. Ilyin, A. V., and V. D. Ilyin. 2015. E-trade with direct lending and normalized money. *Agris on-line Papers in Economics and Informatics* 7(4):57–64.

Received March 11, 2021

Contributors

Ilyin Alexander V. (b. 1975)— Candidate of Science (PhD) in technology, leading scientist, 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, A. A. Dorodnicyn Computing Center, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 40 Vavilov Str., Moscow 119333, Russian Federation; vdilyin@yandex.ru

МАТЕМАТИЧЕСКОЕ И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ АГРОЦЕНОЗА ХЛОПЧАТНИКА С УЧЕТОМ ВОЗРАСТНОЙ СТРУКТУРЫ И С ПРОИЗВОЛЬНЫМИ ТРОФИЧЕСКИМИ ФУНКЦИЯМИ

Р. Н. Одинаев¹, А. Б. Гафоров²

Аннотация: Агроценоз хлопчатника служит примером сложной системы, так как расчленяется на три подсистемы: зооценоз, фитоценоз и микробиоценоз. В свою очередь, каждая подсистема может быть разделена на конечное число более мелких подсистем. Например, зооценоз делится на полезных (хищники и паразиты вредителей) и вредных насекомых (вредителей хлопчатника), фитоценоз состоит из одного элемента — хлопчатника, микробиоценоз состоит из возбудителей вилта и других микроорганизмов. В данной работе рассматривается математическое и компьютерное моделирование процесса защиты агроценоза хлопчатника с учетом возрастной структуры и с произвольными трофическими функциями.

Ключевые слова: математическое и компьютерное моделирование; модель; агроценоз; хлопчатник; произвольная трофическая функция; вредные насекомые; полезные насекомые; возрастная структура; биомасса; численность биологических популяций; нектароносные растения

DOI: 10.14357/08696527210216

1 Введение

Математическое моделирование динамики численности биологических популяций имеет достаточно длительную историю [1–4].

Проблема описания процесса защиты сельскохозяйственного урожая при математическом моделировании остается актуальной. Сегодня проблема защиты урожая привлекает внимание большого круга ученых в связи с перспективным использованием в приоритетных направлениях развития науки, технологии и сельского хозяйства [5–7]. Одной из важнейших народнохозяйственных, социальных и природоохранных проблем в настоящее время стало совершенствование систем защиты сельскохозяйственных культур от вредителей. Исследования устойчивости математических моделей с учетом временной и пространственной связи для модельных популяций, сообществ экосистем изучены в работах Ю. М. Свирежева, Д. О. Логофета, М. К. Юнусова, Р. Н. Одинаева [8–11].

¹Таджикский национальный университет, Душанбе, raim_odinaev@mail.ru

²Таджикский национальный университет, Душанбе, alisher_gaforov@mail.ru

2 Математическая модель агроценоза хлопчатника с учетом возрастной структуры и с произвольными трофическими функциями

В этом разделе рассмотрим математическую модель агроценоза хлопчатника. По результатам исследования ученых Научного центра защиты растений Академии сельскохозяйственных наук Республики Таджикистан, в настоящее время на хлопковых полях наиболее часто встречаются три вида вредителей хлопчатника, т. е. насекомых, которые непосредственно питаются соками и зеленой массой хлопчатника, а также три вида хищников, т. е. полезных насекомых, которые питаются вредителями.

Рассмотрим математическую модель агроценоза, имеющую три трофических уровня: хлопчатник – вредные насекомые – полезные насекомые.

Будем считать, что состояние модельного агроценоза описывается при помощи следующей системы дифференциальных уравнений [9, 10, 12, 13]:

$$\frac{dN_0}{dt} = Q - \alpha_0 N_0 N_1; \quad (1)$$

$$\frac{dN_1}{dt} = k_0 \alpha_0 N_0 N_1 - V_1(N_1) \sum_{i=1}^3 \tilde{N}_{i+1} - m_1 N_1; \quad (2)$$

$$\frac{\partial N_{i-1}}{\partial t} + \frac{\partial N_{i-1}}{\partial a} = k_{i-2} V_1(N_1) N_{i-1} - V_{i-1}(N_{i-1}) \tilde{N}_{i+2} + \delta N_{i+2} - m_{i-1} N_{i-1}, \quad i = \overline{3, 5}; \quad (3)$$

$$\frac{\partial N_{i+2}}{\partial t} + \frac{\partial N_{i+2}}{\partial a} = k_{i+1} V_{i-1}(N_{i-1}) N_{i+2} - \varepsilon N_{i+2}^2 - m_{i+2} N_{i+2}, \quad i = \overline{3, 5}; \quad (4)$$

$$N_i|_{t=0} = N_i^0, \quad i = \overline{0, 7}, \quad N_i(0, t) = \int_0^\infty B_i(\xi) N_i(\xi, t) d\xi, \quad i = \overline{2, 7};$$

$$\tilde{N}_i = \int_{\alpha_i}^{\beta_i} N_i(a, t) da, \quad i = \overline{2, 7},$$

где $N_0 = N_0(t)$ — масса внешнего ресурса в момент времени t (удобрения или вода, используемые для полива, или солнечная энергия); $N_1 = N_1(t)$ — биомасса хлопчатника в момент времени t ; Q — скорость поступления внешнего ресурса; $N_i = N_i(a, t)$, $i = \overline{2, 7}$, — биомасса, или численность, вредных ($i = 2, 3, 4$) и полезных ($i = 5, 6, 7$) насекомых возраста a в момент времени t (N_2 — численность

тли; N_3 — численность хлопковой совки; N_4 — численность паутинного клеща; N_5 — численность трихограммы; N_6 — численность габробракона; N_7 — численность златоглазки); $B_i(\cdot)$ — коэффициент рождаемости вредных и полезных насекомых, $i = \overline{2, 7}$; $V_i(\cdot)$ — трофическая функция со свойствами $V_i(\cdot) \geq 0$, $dV_i/dN > 0$, $d^2V_i/dN^2 \leq 0$, $i = \overline{1, 4}$; δ — биомасса нектароносных растений (например, укроп, хну, базилик сеют в поле хлопчатника полосами на определенном расстоянии, и эти растения в период вегетации привлекают к себе полезных насекомых, а также служат для них источником дополнительной углеродной пищи); α_0 , m_i , k_i и ε — биологические параметры популяций, входящих в агроценоз (m_i — усредненный коэффициент естественной смертности, $i = \overline{1, 2, 3}$; k_i — доля потребленной биомассы, идущая на репродуктивный обмен и рост; α_0 — коэффициент трофических функций; ε — коэффициент самолимитирования популяции полезных насекомых); $\tilde{N}_i = \tilde{N}_i(t)$ — суммарная численность вредных или полезных насекомых в том возрасте, когда они соответственно вредят сельхозкультуре или уничтожают ее вредителей:

$$\tilde{N}_i = \tilde{N}_i(t) = \int_{\alpha_i}^{\beta_i} N_i(a, t) da, \quad \alpha, \beta = const > 0, \quad i = \overline{2, 7};$$

t — время, $t \in [0, \tau]$; $\tau = const < \infty$; a — возраст, $a \in [0, \infty)$.

Предположим, что $B_i(\cdot) \geq 0$, $i = \overline{2, 7}$.

Определение 1. Средней биомассой растения, или средним урожаем, в момент времени τ назовем величину

$$N_1^\tau = \frac{1}{\tau} \int_0^\tau N_1(t) dt.$$

Сформулируем постановку задачи. Пусть

$$\tilde{N}_{i-1}(t) = \frac{1}{\tau} \int_0^\tau \tilde{N}_{i-1}(t) dt, \quad \tilde{N}_{i+2}(t) = \frac{1}{\tau} \int_0^\tau \tilde{N}_{i+2}(t) dt, \quad i = \overline{3, 5}, \quad \tau > 0.$$

Требуется найти числа N_{i-1}^p , $i = \overline{3, 5}$, и N_{i+2}^p , $i = \overline{3, 5}$, такие, что

$$\frac{1}{\tau} \int_0^\tau \tilde{N}_{i-1}(t) dt \leq N_{i-1}^p, \quad i = \overline{3, 5}; \quad \frac{1}{\tau} \int_0^\tau \tilde{N}_{i+2}(t) dt \geq N_{i+2}^p, \quad i = \overline{3, 5},$$

и гарантируется выполнение условия

$$\frac{1}{\tau} \int_0^{\tau} N_1(t) dt \geq N_1^p, \quad N_1^p \in [N_1^{\min}, N_1^{\max}],$$

где $N_1^{\min}, N_1^{\max} = \text{const} > 0$; N_1^p — заданная величина планируемого урожая биомассы хлопчатника.

Для хлопкового агроценоза задача определения порога вредоносности вредителей и уровня эффективности полезных насекомых (энтомофагов) имеет большое практическое значение. Если численность вредных насекомых не превышает порогового уровня, а численность энтомофагов не ниже уровня эффективности, то в агроценозе наступает состояние, при котором нет необходимости в применении химических средств против вредителей.

Определение 2. Величину N_{i-1}^p , $i = \overline{3, 5}$, назовем порогом вредоносности вредных насекомых, а N_{i+2}^p , $i = \overline{3, 5}$, — уровнем эффективности полезных насекомых (энтомофагов).

Основным результатом решения задачи защиты агроценоза хлопчатника с учетом возрастного состава и с произвольными трофическими функциями служит следующее утверждение.

Теорема. *Чтобы имело место условие*

$$\frac{1}{\tau} \int_0^{\tau} N_1(t) dt \geq N_1^p, \quad N_1^p \in [N_1^{\min}, N_1^{\max}],$$

при

$$V_i(\cdot) \geq 0, \quad \frac{dV_i}{dN} > 0, \quad \frac{d^2V_i}{dN^2} \leq 0$$

и

$$0 < \min_{\substack{0 \leq a < \infty \\ 0 \leq t \leq \tau}} \frac{V_1(N_1(t))}{N_1(t)} = \bar{\alpha}_1 < \infty, \quad 0 < \max_{\substack{0 \leq a < \infty \\ 0 \leq t \leq \tau}} \frac{V_{i-1}(N_{i-1}(a, t))}{N_{i-1}(a, t)} = \bar{\alpha}_{i-1} < \infty, \\ i = \overline{3, 5},$$

$\bar{\alpha}_1, \bar{\alpha}_2, \bar{\alpha}_3, \bar{\alpha}_4 = const > 0$, необходимо и достаточно выполнение неравенств:

$$\left. \begin{aligned} N_0(t) &\leq \frac{Q}{\alpha_0 N_1^p}, & 0 \leq t \leq \tau; \\ \frac{1}{\tau} \int_0^\tau \tilde{N}_{i-1}(t) dt &\leq N_{i-1}^p, & i = \overline{3, 5}; \\ \frac{1}{\tau} \int_0^\tau \tilde{N}_{i+2}(t) dt &\geq N_{i+2}^p, & i = \overline{3, 5}, \end{aligned} \right\} \quad (5)$$

где N_{i-1}^p и N_{i+2}^p , $i = \overline{3, 5}$, определяются по следующим формулам:

$$\begin{aligned} N_2^p &= \frac{k_0 Q}{\bar{\alpha}_1 N_1^p} - \frac{m_1}{\bar{\alpha}_1} - \frac{1}{\bar{\alpha}_1 \tau} \ln \frac{N_1(\tau)}{N_1(0)} - \tilde{N}_3(t) - \tilde{N}_4(t); \\ N_3^p &= \frac{k_0 Q}{\bar{\alpha}_1 N_1^p} - \frac{m_1}{\bar{\alpha}_1} - \frac{1}{\bar{\alpha}_1 \tau} \ln \frac{N_1(\tau)}{N_1(0)} - \tilde{N}_2(t) - \tilde{N}_4(t); \\ N_4^p &= \frac{k_0 Q}{\bar{\alpha}_1 N_1^p} - \frac{m_1}{\bar{\alpha}_1} - \frac{1}{\bar{\alpha}_1 \tau} \ln \frac{N_1(\tau)}{N_1(0)} - \tilde{N}_2(t) - \tilde{N}_4(t); \\ N_5^p &= \frac{k_1 \bar{\alpha}_1}{\bar{\alpha}_2} N_1^p + \frac{\delta N_5}{N_2 \bar{\alpha}_2} - \frac{m_2}{\bar{\alpha}_2} - \frac{1}{\tau \bar{\alpha}_2} \max_a \ln \frac{N_2(a, \tau)}{N_2(a, 0)}; \\ N_6^p &= \frac{k_2 \bar{\alpha}_1}{\bar{\alpha}_3} N_1^p + \frac{\delta N_6}{N_3 \bar{\alpha}_3} - \frac{m_3}{\bar{\alpha}_3} - \frac{1}{\tau \bar{\alpha}_3} \max_a \ln \frac{N_3(a, \tau)}{N_3(a, 0)}; \\ N_7^p &= \frac{k_3 \bar{\alpha}_1}{\bar{\alpha}_4} N_1^p + \frac{\delta N_7}{N_4 \bar{\alpha}_4} - \frac{m_4}{\bar{\alpha}_4} - \frac{1}{\tau \bar{\alpha}_4} \max_a \ln \frac{N_4(a, \tau)}{N_4(a, 0)}. \end{aligned}$$

Доказательство. Пусть выполняются условия теоремы.

В силу уравнения (1) получаем:

$$\begin{aligned} N_0(t) &= N_0(0) \exp(-\alpha_0 t) + Q \int_0^\tau \exp(-\alpha_0(t-\tau)) \int_\tau^\zeta N_1(\zeta) d\tau d\zeta \leq \\ &\leq \left[N_0(0) - \frac{Q}{\alpha_0 N_1^p} \right] \exp(-\alpha_0 N_1^p t) + \frac{Q}{\alpha_0 N_1^p} = \frac{Q}{\alpha_0 N_1^p}, \quad 0 \leq t \leq \tau, \end{aligned}$$

где

$$N_0(0) = \frac{Q}{\alpha_0 N_1^p}.$$

В силу уравнения (2) имеем:

$$\begin{aligned} \frac{dN_1}{dt} &= k_0\alpha_0 N_0 N_1 - \sum_{i=1}^3 V_1(N_1)\tilde{N}_{i+1} - m_1 N_1; \\ \frac{d(\ln N_1)}{dt} &= k_0\alpha_0 N_0 - \frac{V_1(N_1)}{N_1} \sum_{i=1}^3 \tilde{N}_{i+1} - m_1; \\ \frac{V_1(N_1)}{N_1} \sum_{i=1}^3 \tilde{N}_{i+1}(t) &= \frac{k_0 Q}{N_1^p} - m_1 - \frac{d(\ln N_1)}{dt}. \end{aligned} \quad (6)$$

Интегрируя равенство (6) по t от 0 до τ , получаем

$$\frac{1}{\tau} \int_0^\tau \sum_{i=1}^3 \tilde{N}_{i+1}(t) dt \leq \frac{k_0 Q}{\alpha_1 N_1^p} - \frac{m_a}{\alpha_1} - \frac{1}{\alpha_1 \tau} \ln \frac{N_1(\tau)}{N_1(0)} = N_{i+1}^p,$$

отсюда

$$\frac{1}{\tau} \int_0^\tau \tilde{N}_{i-1}(t) dt \leq N_{i-1}^p, \quad i = \overline{3, 5}.$$

На основе уравнения (3) имеем:

$$\begin{aligned} \frac{\partial N_{i-1}}{\partial t} + \frac{\partial N_{i-1}}{\partial a} &= k_{i-2} V_1(N_1) N_{i-1} - V_{i-1}(N_{i-1}) \tilde{N}_{i+2} - \delta N_{i+2} - m_{i-1} N_{i-1}, \\ & i = \overline{3, 5}. \end{aligned}$$

Введем обозначения $a = t + \xi$; $\varphi(t, \xi) = N_{i-1}(a, t)$, $i = \overline{3, 5}$. Тогда

$$\frac{\partial \varphi}{\partial t} = \frac{\partial N_{i-1}}{\partial t} + \frac{\partial N_{i-1}}{\partial a}, \quad i = \overline{3, 5}.$$

Получаем

$$\begin{aligned} \frac{\partial}{\partial t} \ln \varphi &= k_{i-2} V_1(N_1) - \frac{V_{i-1}(N_{i-1}) \tilde{N}_{i+2}}{N_{i-1}} - \frac{\delta N_{i+2}}{N_{i-1}} - m_{i-1}, \quad i = \overline{3, 5}, \\ \frac{V_{i-1}(N_{i-1})}{N_{i-1}} \tilde{N}_{i+2} &= k_{i-2} V_1(N_1) - \frac{\delta N_{i+2}}{N_{i-1}} - m_{i-1} - \frac{\partial}{\partial t} \ln \varphi, \quad i = \overline{3, 5}. \end{aligned} \quad (7)$$

Интегрируя уравнение (7) по t от 0 до τ , получаем

$$\begin{aligned} \frac{1}{\tau} \int_0^{\tau} \frac{V_{i-1}(N_{i-1})}{N_{i-1}} \tilde{N}_{i+2}(t) dt &= \\ &= k_{i-2} \bar{\alpha}_1 N_1^p - \frac{\delta N_{i+2}}{N_{i-1}} - m_{i-1} - \frac{1}{\tau} \ln \frac{N_{i-1}(a, \tau)}{N_{i-1}(a, 0)}, \quad i = \overline{3, 5}, \end{aligned}$$

и, следовательно,

$$\frac{1}{\tau} \int_0^{\tau} \tilde{N}_{i+2}(t) dt \geq N_{i+2}^p, \quad i = \overline{3, 5}.$$

Теперь оценим $\max N_{i+2}(t)$, $i = \overline{3, 5}$.

Из уравнения (4) получаем

$$N_{i+2}(t) = \frac{N_{i+2}(0) \exp \int_0^t A(\xi) d\xi}{1 + \varepsilon N_{i+2}(0) \int_0^t \exp \left(\int_{\tau}^t A(\xi) d\xi \right) d\tau} \leq N_{\max}, \quad N_{\max} = \text{const} < \infty,$$

где $A(\xi) = k_{i+1} V_{i-1}(N_{i-1}) - m_{i+2}$.

Достаточность. Пусть выполняются неравенства (5). Докажем, что имеет место условие

$$\frac{1}{\tau} \int_0^{\tau} N_1(t) dt \geq N_1^p.$$

Из уравнения (1) находим

$$N_0(t) = N_0(0) + Qt - \alpha_0 \int_0^{\tau} N_0 N_1(t) dt,$$

отсюда

$$N_0(t) \geq N_0(0) + Qt - \frac{Q}{N_1^p} \int_0^{\tau} N_1(t) dt$$

и, следовательно,

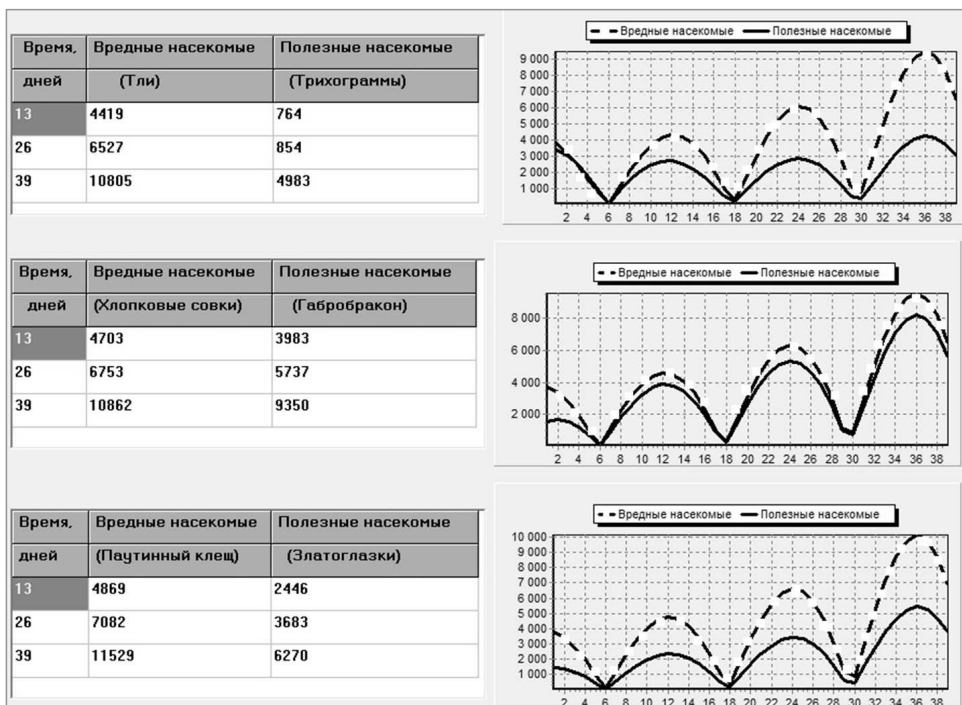
$$\frac{1}{\tau} \int_0^{\tau} N_1(t) dt - N_1^p \geq \frac{(N_0(0) - Q/(\alpha_0 N_1^p)) N_1^p}{Qt},$$

откуда вытекает

$$\frac{1}{\tau} \int_0^{\tau} N_1(t) dt \geq N_1^p.$$

3 Результаты компьютерного моделирования

В данном разделе приведены результаты компьютерного моделирования. Компьютерная программа разработана на языке программирования С++, широко используемом в научных, инженерных, математических и компьютерных областях. Программа предназначена для вычисления динамики численности вредных и полезных насекомых и дает возможность прогнозировать динамику развития насекомых на рассматриваемый интервал времени. Результаты работы программы можно выводить в виде таблицы и в виде графика. Вычисления проводились в нескольких вариантах на основе предварительных данных. Каждый



Скриншот результатов, полученных при работе программы «Динамика численности развития вредных и полезных насекомых хлопкового агроценоза»

вариант выполнялся при различных начальных значениях численности насекомых и разных интервалах времени. Продолжительность жизни одного поколения принималась равной 13 дням, а число поколений за сезон — 3.

Полученные результаты приведены на рисунке.

4 Заключение

Результаты исследования могут быть использованы для решения задачи прогнозирования и планирования, проведения натурных экспериментов для конкретных популяций, биологических сообществ и экологических систем.

Литература

1. *Вольтерра В.* Математическая теория борьбы за существование / Пер. с фр. — М.: Наука, 1976. 286 с. (*Volterra V.* Leçons sur la théorie mathématique de la lutte pour la vie. — Paris, 1931. 222 p.)
2. *Lotka A. T.* Theorie analytique des associations biologiques. — Paris: Principles, 1939. 218 p.
3. *Lotka A. T.* Elements of physical biology. — New York, NY, USA: Dover Publ., 1956. 460 p.
4. *May R. M.* Simple mathematical model with very complicated dynamics // *Nature*, 1976. Vol. 261. Iss. 560. P. 459–467.
5. *Мурри Дж.* Нелинейные дифференциальные уравнения в биологии / Пер. с англ. — М.: Мир, 1983. 397 с. (*Murray J. D.* Lectures on nonlinear-differential-equation models in biology. — Clarendon Press, 1977. 370 p.)
6. *Liu B., Zhang Y., Chen I.* The dynamical behaviors of Lotka–Volterra predator-prey model concerning integrated pest management // *Nonlinear Anal. — Real*, 2005. Vol. 6. Iss. 2. P. 227–243. doi: 10.1016/j.nonrwa.2004.08.001.
7. *Jambhulkar P. P., Murlidhar S.* Information technology in plant protection // *Current concepts in plant protection / Ed. S. Banik.* — New Dehli: Studium Press India Pvt. Ltd., 2013. P. 30.
8. *Свирижев Ю. М., Логофет Д. О.* Устойчивость биологических сообществ. — М.: Наука, 1978. 352 с.
9. *Логофет Д. О., Юнусов М. К.* Вопросы качественной устойчивости и регуляризации в динамических моделях агробиоценоза хлопчатника // *Вопросы кибернетики*, 1979. Вып. 52. С. 60–72.
10. *Одинаев Р. Н.* Исследование математической модели задачи защиты растений в стационарном случае // *Вестник Таджикского национального университета. Сер. естественных наук*, 2013. № 1/3(110). С. 7–11.
11. *Одинаев Р. Н.* Математическая модель задачи защиты растений в биосистеме типа «вредные насекомые – полезные насекомые» с произвольными трофическими функциями // *Системы и средства информатики*, 2019. Т. 29. Вып. 1. С. 96–108. doi: 10.14357/08696527190109.

12. *Одинаев Р. Н.* Численный метод решения интегро-дифференциальной задачи защиты растений // Вестник Таджикского национального университета. Сер. естественных наук, 2017. № 1-5. С. 112–116.
13. *Одинаев Р. Н.* Компьютерное моделирование процесса защиты растений с учетом временно-возрастной структуры и пространственного распределения с произвольными трофическими функциями // Наука, новые технологии и инновации Кыргызстана. — Бишкек, 2018. Вып. 2. С. 33–38.

Поступила в редакцию 01.07.20

MATHEMATICAL AND COMPUTER MODELING OF AGROCENOSIS OF COTTON TAKING INTO ACCOUNT THE AGE STRUCTURE AND WITH ARBITRARY TROPHIC FUNCTIONS

R. N. Odinaev and A. B. Gaforov

Faculty of Mechanics and Mathematics, Tajik National University, 17 Rudaki Av.,
Dushanbe 734025, Republic of Tajikistan

Abstract: Cotton agrocenosis is an example of a complex system, since it is divided into three subsystems (zoocenosis, phytocenosis, and microbiocenosis), and, in turn, each subsystem can be divided into a finite number of smaller subsystems. For example, zoocenosis is divided into beneficial (predators and parasites of pests) and harmful insects (cotton pests), phytocenosis consists of one element of cotton, microbiocenosis consists of pathogens of wilt and other microorganisms. This paper considers mathematical and computer modeling of the process of protecting agrocenosis of cotton taking into account the age structure and arbitrary trophic functions.

Keywords: mathematical and computer modeling; model; agrocenosis; cotton; arbitrary trophic function; harmful insects; beneficial insects; age structure; biomass; number of biological populations; nectariferous plants

DOI: 10.14357/08696527210216

References

1. Volterra, V. 1931. *Leçons sur la théorie mathématique de la lutte pour la vie*. Paris. 222 p.
2. Lotka, A. T. 1939. *Theoric analytique des associations biologiques*. Paris: Principes. 218 p.
3. Lotka, A. T. 1956. *Elements of physical biology*. New York, NY: Dover Publ. 460 p.
4. May, R. M. 1976. Simple mathematical model with very complicated dynamics. *Nature* 261(560):459–467.

5. Murray, J. D. 1977. *Lectures on nonlinear-differential-equation models in biology*. Clarendon Press. 370 p.
6. Liu, B., Y. Zhang, and I. Chen. 2005. The dynamical behaviors of Lotka–Volterra predator–prey model concerning integrated pest management. *Nonlinear Anal. — Real* 6(2):227–243. doi: 10.1016/j.nonrwa.2004.08.001.
7. Jambhulkar, P. P., and S. Murlidhar. 2013. Information technology in plant protection. *Current concepts in plant protection*. Ed. S. Banik. New Dehli: Studium Press India Pvt. Ltd. 30.
8. Svirezhev, Yu. M., and D. O. Logofet. 1978. *Ustoychivost' biologicheskikh soobshchestv* [Stability of biological communities]. Moscow: Nauka. 352 p.
9. Logofet, D. O., and M. K. Yunusov. 1979. Voprosy kachestvennoy ustoychivosti i regulyarizatsii v dinamicheskikh modelyakh agrobiotsenoza khlopchatnika [Questions of qualitative stability and regularization in dynamic models of cotton agrophytocenosis]. *Voprosy kibernetiki* [Problems of Cybernetics] 52:60–72.
10. Odinaev, R. N. 2013. Issledovanie matematicheskoy modeli zadachi zashchity rasteniy v statsionarnom sluchae [The study of the mathematical model of the plant protection problem in a stationary case]. *Vestnik Tadzhikskogo natsional'nogo universiteta. Ser. estestvennykh nauk* [Bull. Tajik National University. Ser. Natural Sciences] 1/3(110):7–11.
11. Odinaev, R. N. 2019. Matematicheskaya model' zadachi zashchity rasteniy v biosisteme tipa “vrednye nasekomye – poleznye nasekomye” s proizvol'nymi troficheskimi funktsiyami [A mathematical model of the problem of plant protection in a biosystem of the type “harmful insects – beneficial insects” with arbitrary trophic functions]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(1):96–108. doi: 10.14357/08696527190109.
12. Odinaev, R. N. 2017. Chislenny metod resheniya integro-differentsial'noy zadachi zashchity rasteniy [A numerical method for solving the integrodifferential problem of plant protection]. *Vestnik Tadzhikskogo natsional'nogo universiteta. Ser. estestvennykh nauk* [Bull. Tajik National University. Ser. Natural Sciences] 1-5:112–116.
13. Odinaev, R. N. 2018. Komp'yuternoe modelirovanie protsessa zashchity rasteniy s uchetom vremenno-voznrastnoy struktury i prostranstvennogo raspredeleniya s proizvol'nymi troficheskimi funktsiyami [Computer simulation of the plant protection process taking into account the temporal age structure and spatial distribution with arbitrary trophic functions]. *Nauka, novye tekhnologii i innovatsii Kyrgyzstana* [Science, New Technologies, and Innovations of Kyrgyzstan] 2:33–38.

Received July 1, 2020

Contributors

Odinaev Raim N. (b. 1964) — Doctor of Science in physics and mathematics, associate professor, dean, Faculty of Mechanics and Mathematics, Tajik National University, 17 Rudaki Av., Dushanbe 734025, Republic of Tajikistan; aim.odinaev@mail.ru

Gaforov Alisher B. (b. 1995) — PhD student, Department of Mathematical and Computer Modeling, Tajik National University, 17 Rudaki Av., Dushanbe 734025, Republic of Tajikistan; alisher_gaforov@mail.ru

**ПОПРАВКА К СТАТЬЕ О. М. КОРЧАЖКИНОЙ «SIR-МОДЕЛЬ
КАК ИНСТРУМЕНТ ИССЛЕДОВАНИЯ ДЕСТРУКТИВНЫХ
ПРОЦЕССОВ ПРИ УСВОЕНИИ НОВОГО ЗНАНИЯ»
(Системы и средства информатики, 2021. Т. 31. № 1. С. 168–180)
DOI: 10.14357/08696527210114**

С. 173, формула (5): вместо

$$\frac{\partial I}{\partial t} = \dot{I} - \beta SI - \gamma I; \quad (5)$$

должно быть

$$\frac{\partial I}{\partial t} = \dot{I} = \beta SI - \gamma I; \quad (5)$$

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 Павел Олегович (р. 1979) — кандидат технических наук, директор Орловского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

Борисов Андрей Владимирович (р. 1965) — доктор физико-математических наук, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Босов Алексей Вячеславович (р. 1969) — доктор технических наук, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Вакуленко Василий Васильевич (р. 1995) — инженер-исследователь 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

Воронцов Михаил Олегович (р. 1996) — студент факультета вычислительной математики и кибернетики Московского государственного университета имени М. В. Ломоносова

Гайдамака Юлия Васильевна (р. 1971) — доктор физико-математических наук, профессор кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов; старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Гаранин Александр Иванович (р. 1951) — кандидат технических наук, старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Гафоров Алишер Бобобекович (р. 1995) — аспирант кафедры математического и компьютерного моделирования Таджикского национального университета

Грушо Александр Александрович (р. 1946) — доктор физико-математических наук, профессор, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Грушо Николай Александрович (р. 1982) — кандидат физико-математических наук, старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Денисов Сергей Анатольевич (р. 1985) — ведущий инженер Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Егоров Владимир Борисович (р. 1948) — кандидат технических наук, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Егорова Анна Юрьевна (р. 1991) — младший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Жуков Денис Владимирович (р. 1979) — главный специалист Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Забейайло Михаил Иванович (р. 1956) — доктор физико-математических наук, доцент, главный научный сотрудник Вычислительного центра им. А. А. Дородницына Федерального исследовательского центра «Информатика и управление» Российской академии наук

Зацаринный Александр Алексеевич (р. 1951) — доктор технических наук, профессор, заместитель директора Федерального исследовательского центра «Информатика и управление» Российской академии наук (ФИЦ ИУ 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

Иванов Алексей Владимирович (р. 1976) — кандидат технических наук, старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Ильин Александр Владимирович (р. 1975) — кандидат технических наук, ведущий научный сотрудник 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

Кириков Игорь Александрович (р. 1955) — кандидат технических наук, директор Калининградского филиала 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 имени 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

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

Нурiev Виталий Александрович (р. 1980) — кандидат филологических наук, старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Одинаев Раим Назарович (р. 1964) — доктор физико-математических наук, доцент, декан механико-математического факультета Таджикского национального университета

Ступников Сергей Александрович (р. 1978) — кандидат 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

Хайров Эмиль Маратович (р. 1997) — студент магистратуры кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов

Царев Алексей Сергеевич (р. 1994) — аспирант кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов

Шанин Иван Андреевич (р. 1991) — младший 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

Яковлев Олег Альбертович (р. 1992) — младший научный сотрудник Орловского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

Правила подготовки рукописей статей для публикации в журнале «Системы и средства информатики»

Журнал «Системы и средства информатики» публикует теоретические, обзорные и дискуссионные статьи, посвященные научным исследованиям и разработкам в области информационных технологий.

Журнал издается на русском языке. По специальному решению редколлегии отдельные статьи могут печататься на английском языке.

Тематика журнала охватывает следующие направления:

- информационно-телекоммуникационные системы и средства их построения;
- архитектура и программное обеспечение вычислительных машин, комплексов и сетей;
- методы и средства защиты информации.

1. В журнале печатаются статьи, содержащие результаты, ранее не опубликованные и не предназначенные к одновременной публикации в других изданиях.

Публикация предоставленной автором(ами) рукописи не должна нарушать положения глав 69, 70 раздела VII части IV Гражданского кодекса, которые определяют права на результаты интеллектуальной деятельности и средства индивидуализации, в том числе авторские права, в РФ.

Ответственность за нарушение авторских прав, в случае предъявления претензий к редакции журнала, несут авторы статей.

Направляя рукопись в редакцию, авторы сохраняют свои права на данную рукопись и при этом передают учредителям и редколлегии журнала неисключительные права на издание статьи на русском языке (или на языке статьи, если он отличен от русского) и на перевод ее на английский язык, а также на ее распространение в России и за рубежом. Каждый автор должен представить в редакцию подписанный с его стороны «Лицензионный договор о передаче неисключительных прав на использование произведения», текст которого размещен по адресу <http://www.ipiran.ru/publications/licence.doc>. Этот договор может быть представлен в бумажном (в 2-х экз.) или в электронном виде (отсканированная копия заполненного и подписанного документа).

Редколлегия вправе запросить у авторов экспертное заключение о возможности публикации представленной статьи в открытой печати.

2. К статье прилагаются данные автора (авторов) (см. п. 8). При наличии нескольких авторов указывается фамилия автора, ответственного за переписку с редакцией.
3. Редакция журнала осуществляет экспертизу присланных статей в соответствии с принятой в журнале процедурой рецензирования.

Возвращение рукописи на доработку не означает ее принятия к печати.

Доработанный вариант с ответом на замечания рецензента необходимо прислать в редакцию.

4. Решение редколлегии о публикации статьи или ее отклонении сообщается авторам. Редколлегия может также направить авторам текст рецензии на их статью. Дискуссия по поводу отклоненных статей не ведется.

5. Редактура статей высылается авторам для просмотра. Замечания к редакции должны быть присланы авторами в кратчайшие сроки.
6. Рукопись предоставляется в электронном виде в форматах MS WORD (.doc или .docx) или L^AT_EX (.tex), дополнительно — в формате .pdf, на дискете, лазерном диске или электронной почтой. Предоставление бумажной рукописи необязательно.
7. При подготовке рукописи в MS Word рекомендуется использовать следующие настройки.

Параметры страницы: формат — А4; ориентация — книжная; поля (см): внутри — 2,5, снаружи — 1,5, сверху и снизу — 2, от края до нижнего колонтитула — 1,3.

Основной текст: стиль — «Обычный», шрифт — Times New Roman, размер — 14 пунктов, абзацный отступ — 0,5 см, 1,5 интервала, выравнивание — по ширине.

Рекомендуемый объем рукописи — не свыше 10 страниц указанного формата. При превышении указанного объема редколлегия вправе потребовать от автора сокращения объема рукописи.

Сокращения слов, помимо стандартных, не допускаются. Допускается минимальное количество аббревиатур.

Все страницы рукописи нумеруются.

Шаблоны примеров оформления представлены в Интернете:

<http://www.ipiran.ru/publications/collected/template.doc>

8. Статья должна содержать следующую информацию на **русском и английском языках**:
 - название статьи;
 - Ф.И.О. авторов, на английском можно только имя и фамилию;
 - место работы, с указанием города и страны и электронного адреса каждого автора;
 - сведения об авторах, в соответствии с форматом, образцы которого представлены на страницах:
http://www.ipiran.ru/journal/collected/2019_29_03_rus/authors.asp и
http://www.ipiran.ru/journal/collected/2019_29_03_eng/authors.asp;
 - аннотация (не менее 100 слов на каждом из языков). Аннотация — это краткое резюме работы, которое может публиковаться отдельно. Она является основным источником информации в информационных системах и базах данных. Английская аннотация должна быть оригинальной, может не быть дословным переводом русского текста и должна быть написана хорошим английским языком. В аннотации не должно быть ссылок на литературу и, по возможности, формул;
 - ключевые слова — желательно из принятых в мировой научно-технической литературе тематических тезаурусов. Предложения не могут быть ключевыми словами.
 - источники финансирования работы (ссылка на гранты, проекты, поддерживающие организации и т. п.).
9. Требования к спискам литературы.

Ссылки на литературу в тексте статьи нумеруются (в квадратных скобках) и располагаются в каждом из списков литературы в порядке первых упоминаний.

Списки литературы представляются в двух вариантах:

- (1) **Список литературы к русскоязычной части.** Русские и английские работы — на языке и в алфавите оригинала.
- (2) **References.** Русские работы и работы на других языках — в латинской транслитерации с переводом на английский язык; английские работы и работы на других языках — на языке оригинала.

Необходимо для составления списка “References” пользоваться размещенной на сайте <http://www.translit.net/ru/bgn/> бесплатной программой транслитерации русского текста в латиницу.

Список литературы “References” приводится полностью отдельным блоком, повторяя все позиции из списка литературы к русскоязычной части, независимо от того, имеются или нет в нем иностранные источники. Если в списке литературы к русскоязычной части есть ссылки на иностранные публикации, набранные латиницей, они полностью повторяются в списке “References”.

Примеры ссылок на различные виды публикаций в списке “References”:

Описание статьи из журнала:

Zhang, Z., and D. Zhu. 2008. Experimental research on the localized electrochemical micromachining. *Russ. J. Electrochem.* 44(8):926–930. doi:10.1134/S1023193508080077.

Описание статьи из электронного журнала:

Swaminathan, V., E. Lepkoswka-White, and B. P. Rao. 1999. Browsers or buyers in cyberspace? An investigation of electronic factors influencing electronic exchange. *JCMC* 5(2). Available at: <http://www.ascusc.org/jcmc/vol5/issue2/> (accessed April 28, 2011).

Описание материалов конференций:

Usmanov, T. S., A. A. Gusmanov, I. Z. Mullagalin, R. Ju. Muhametshina, A. N. Chervyakova, and A. V. Sveshnikov. 2007. Osobennosti proektirovaniya razrabotki mestorozhdeniy s primeneniem gidrorazryva plasta [Features of the design of field development with the use of hydraulic fracturing]. *Trudy 6-go Mezhdunarodnogo Simpoziuma “Novye resursoberegayushchie tekhnologii nedropol’zovaniya i povyshe-niya neftegazootdachi”* [6th Symposium (International) “New Energy Saving Subsoil Technologies and the Increasing of the Oil and Gas Impact” Proceedings]. Moscow. 267–272.

Описание книги (монографии, сборники):

Lindorf, L. S., and L. G. Mamikonians, eds. 1972. *Ekspluatatsiya turbogeneratorov s neposredstvennym okhlazhdeniem* [Operation of turbine generators with direct cooling]. Moscow: Energy 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. — New York, NY, USA: Wiley, 1974. 521 p.)

2. В англоязычной части:

Timoshenko, S. P., D. H. Young, and W. Weaver. 1974. *Vibration problems in engineering*. 4th ed. New York, NY: Wiley. 521 p.

Описание неопубликованного документа:

Latypov, A. R., M. M. Khasanov, and V. A. Baikov. 2004. Geology and production (NGT GiD). Certificate on official registration of the computer program No. 2004611198. (In Russian, unpubl.)

Описание интернет-ресурса:

Pravila tsitirovaniya istochnikov [Rules for the citing of sources]. Available at: <http://www.scribd.com/doc/1034528/> (accessed February 7, 2011).

Описание диссертации или автореферата диссертации:

Semenov, V. I. 2003. Matematicheskoe modelirovanie plazmy v sisteme kompaktny tor [Mathematical modeling of the plasma in the compact torus]. Moscow. D.Sc. Diss. 272 p.

Kozhunova, O. S. 2009. Tekhnologiya razrabotki semanticheskogo slovary informatzionnogo monitoringa [Technology of development of semantic dictionary of information monitoring system]. PhD Thesis. Moscow: IPI RAN. 23 p.

Описание ГОСТа:

GOST 8.586.5-2005. 2007. Metodika vypolneniya izmereniy. Izmerenie raskhoda i kolichestva zhidkostey i gazov s pomoshch'yu standartnykh suzhayushchikh ustroystv [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. Moscow: Standardinform Publs. 10 p.

Описание патента:

Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. Spособ orientirovaniya po krenu letatel'nogo apparata s opticheskoy golovkoy samonavedeniya [The way to orient on the roll of aircraft with optical homing head]. Patent RF No. 2280590.

10. Присланные в редакцию материалы авторам не возвращаются.
11. При отправке файлов по электронной почте просим придерживаться следующих правил:
 - указывать в поле subject (тема) название журнала и фамилию автора;
 - использовать attach (присоединение);
 - в состав электронной версии статьи должны входить: файл, содержащий текст статьи, и файл(ы), содержащий(е) иллюстрации.
12. Журнал «Системы и средства информатики» является некоммерческим изданием. Плата за публикацию не взимается, гонорар авторам не выплачивается.

Адрес редакции журнала «Системы и средства информатики»:

Москва 119333, ул. Вавилова, д. 44, корп. 2, ФИЦ ИУ РАН

Тел.: +7 (499) 135-86-92 Факс: +7 (495) 930-45-05

e-mail: ssi@frccsc.ru (Стригина Светлана Николаевна)

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

Requirements for manuscripts submitted to Journal “Systems and Means of Informatics”

Journal “Systems and Means of Informatics” publishes theoretical, review, and discussion articles on the research and development in the field of information technology.

The journal is published in Russian. By a special decision of the editorial board, some articles can be published in English.

Topics covered include the following areas:

- information and communication systems and tools of their design;
- architecture and software of computational complexes and networks; and
- methods and tools of information protection.

1. The Journal publishes original articles which have not been published before and are not intended for simultaneous publication in other editions. An article submitted to the Journal must not violate the Copyright law. Sending the manuscript to the Editorial Board, the authors retain all rights of the owners of the manuscript and transfer the nonexclusive rights to publish the article in Russian (or the language of the article, if not Russian) and its distribution in Russia and abroad to the Founders and the Editorial Board. Authors should submit a letter to the Editorial Board in the following form:

Agreement on the transfer of rights to publish:

“We, the undersigned authors of the manuscript “. . . ,” pass to the Founder and the Editorial Board of the Journal “Systems and Means of Informatics” the nonexclusive right to publish the manuscript of the article in Russian (or in English) in both print and electronic versions of the Journal. We affirm that this publication does not violate the Copyright of other persons or organizations.

Author(s) signature(s): (name(s), address(es), date).”

This agreement should be submitted in paper form or in the form of a scanned copy (signed by the authors).

The Editorial Board has the right to request from the authors an official expert conclusion that the submitted article has no classified data prohibited for publication.

2. A submitted article should be attached with **the data on the author(s)** (see item 8). If there are several authors, the contact person should be indicated who is responsible for correspondence with the Editorial Board and other authors about revisions and final approval of the proofs.
3. The Editorial Board of the Journal examines the article according to the established reviewing procedure. If authors receive their article for correction after reviewing, it does not mean that the article is approved to be published. The corrected article should be sent to the Editorial Board for the subsequent review and approval.
4. The decision on the article publication or its rejection is communicated to the authors. The Editorial Board may also send the reviews on the submitted articles to the authors. Any discussion upon the rejected articles is not possible.
5. The edited articles will be sent to the authors for proofread. The comments of the authors to the edited text of the article should be sent to the Editorial Board as soon as possible.
6. The manuscript of the article should be presented electronically in the MS WORD (.doc or .docx) or L^AT_EX (.tex) formats, and additionally in the .pdf format. All documents

may be sent by e-mail or provided on a CD or diskette. A hard copy submission is not necessary.

7. The recommended typesetting instructions for manuscript.

Pages parameters: format A4, portrait orientation, document margins (cm): left — 2.5, right — 1.5, above — 2.0, below — 2.0, footer 1.3.

Text: font — Times New Roman, font size — 14, paragraph indent — 0.5, line spacing — 1.5, justified alignment.

The recommended manuscript size: not more than 10 pages of the specified format. If the specified size exceeded, the editorial board is entitled to require the author to reduce the manuscript.

Use only standard abbreviations. Avoid abbreviations in the title and abstract. The full term for which an abbreviation stands should precede its first use in the text unless it is a standard unit of measurement.

All pages of the manuscript should be numbered.

The templates for the manuscript typesetting are presented on site:

<http://www.ipiran.ru/publication/collected/template.doc>

8. Articles should enclose data both in **Russian and English**:

- title;
- author's name and surname;
- affiliation — organization, its address with ZIP code, city, country, and official e-mail address;
- data on authors according to the format (see site):
http://www.ipiran.ru/journal/collected/2019_29_03_rus/authors.asp and
http://www.ipiran.ru/journal/collected/2019_29_03_eng/authors.asp;
- abstract (not less than 100 words) both in Russian and in English. Abstract is a short summary of the article that can be published separately. The abstract is the main source of information on the article and it could be included in leading information systems and data bases. The abstract in English has to be an original text and should not be an exact translation of the Russian one. Good English is required. In abstracts, avoid references and formulae.
- Indexing is performed on the basis of keywords. The use of keywords from the internationally accepted thematic Thesauri is recommended.
Important! Keywords must not be sentences.
- Acknowledgments.

9. References. Russian references have to be presented both in English translation and in Latin transliteration (refer <http://www.translit.net/ru/bgn/>).

Please take into account the following examples of Russian references appearance:

Article in journal:

Zhang, Z., and D. Zhu. 2008. Experimental research on the localized electrochemical micromachining. *Russ. J. Electrochem.* 44(8):926–930. doi:10.1134/S1023193508080077.

Journal article in electronic format:

Swaminathan, V., E. Lepkoswka-White, and B. P. Rao. 1999. Browsers or buyers in cyberspace? An investigation of electronic factors influencing electronic

exchange. *JCMC* 5(2). Available at: <http://www.ascusc.org/jcmc/vol5/issue2/> (accessed April 28, 2011).

Conference proceedings:

Usmanov, T. S., A. A. Gusmanov, I. Z. Mullagalin, R. Ju. Muhametshina, A. N. Chervyakova, and A. V. Sveshnikov. 2007. Osobennosti proektirovaniya razrabotki mestorozhdeniy s primeneniem gidrorazryva plasta [Features of the design of field development with the use of hydraulic fracturing]. *Trudy 6-go Mezhdunarodnogo Simpoziuma "Novye resursoberegayushchie tekhnologii nedropol'zovaniya i povyshe-niya neftegazootdachi"* [6th Symposium (International) "New Energy Saving Subsoil Technologies and the Increasing of the Oil and Gas Impact" Proceedings]. Moscow. 267–272.

Books and other monographs:

Lindorf, L. S., and L. G. Mamikonians, eds. 1972. *Ekspluatatsiya turbogenera-torov s neposredstvennym okhlazhdeniem* [Operation of turbine generators with direct cooling]. Moscow: Energy Publs. 352 p.

Dissertation and Thesis:

Kozhunova, O. S. 2009. Tekhnologiya razrabotki semanticheskogo slovaryia informat-sionnogo monitoringa [Technology of development of semantic dictionary of information monitoring system]. Moscow: IPI RAN. PhD Thesis. 23 p.

State standards and patents:

GOST 8.586.5-2005. 2007. Metodika vypolneniya izmereniy. Izmerenie raskhoda i kolichestva zhidkostey i gazov s pomoshch'yu standartnykh suzhayushchikh ustroystv [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. M.: Standardinform Publs. 10 p.

Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. Sposob orientirovaniya po krenu letatel'nogo apparata s opticheskoy golovkoy samonavedeniya [The way to orient on the roll of aircraft with optical homing head]. Patent RF No. 2280590.

References in Latin transcription are presented in the original language.

References in the text are numbered according to the order of their first appearance; the number is placed in square brackets. All items from the reference list should be cited.

10. Manuscripts and additional materials are not returned to Authors by the Editorial Board.
11. Submissions of files by e-mail must include:
 - the journal title and author's name in the "Subject" field;
 - an article and additional materials have to be attached using the "attach" function;
 - an electronic version of the article should contain the file with the text and a separate file with figures.
12. "System and Means of Informatics" journal is not a profit publication. There are no charges for the authors as well as there are no royalties.

Editorial Board address:

FRC CSC RAS, 44, block 2, Vavilov Str., Moscow 119333, Russia

Ph.: +7 (499) 135 86 92, Fax: +7 (495) 930 45 05

e-mail: ssi@frccsc.ru (to Svetlana Strigina)

http://www.ipiran.ru/english/journal_systems.asp

SYSTEMS AND MEANS OF INFORMATICS (SISTEMY I SREDSTVA INFORMATIKI)

SCIENTIFIC JOURNAL

Volume 31 No.2 Year 2021

Editor-in-Chief and Chair of Editorial Council
Academician I. A. Sokolov

IN THIS ISSUE:

HIDDEN IMPACT WITHOUT MALICIOUS CODE <i>A. A. Grusho, N. A. Grusho, M. I. Zabezhailo, and E. E. Timonina</i>	4
ANALYTICAL MODEL OF CARRIER SENSE MULTIPLE ACCESS PROTOCOL FOR INDUSTRIAL INTERNET OF THINGS APPLICATIONS <i>A. E. Tsarev, E. M. Khayrov, Yu. V. Gaidamaka, and S. Ya. Shorgin</i>	16
ENSURING THE RELIABILITY OF THE FRC CSC RAS CENTER FOR COLLECTIVE USE <i>A. A. Zatsarinny, A. I. Garanin, and S. A. Denisov</i>	26
ELECTROENCEPHALOGRAPHY DATA ANALYSIS WITH CONVOLUTIONAL AND RECURRENT NEURAL NETWORKS <i>I. A. Shanin and S. A. Stupnikov</i>	36
STIMULATION OF AGENT CONFLICTS IN HYBRID INTELLIGENT MULTIAGENT SYSTEMS <i>S. V. Listopad and I. A. Kirikov</i>	47
INDOOR EXPLORATION STRATEGY FOR AUTONOMOUS MOBILE ROBOT BASED ON DIRECTION MAP <i>O. A. Yakovlev and P. O. Arkhipov</i>	59
SOME ISSUES OF THE SOFTWARE-DEFINED STORAGE <i>V. B. Egorov</i>	70
INFORMATION ASPECTS OF SECURITY IN TRANSPORT: SEARCH AND SELECTION OF INFORMATION <i>A. V. Borisov, A. V. Bosov, D. V. Zhukov, and A. V. Ivanov</i>	80
RESEARCH AND DEVELOPMENT STRATEGY IN THE FIELD OF ARTIFICIAL INTELLEGENCE II: COMPARATIVE ANALYSIS OF SCIENTOMETRIC INDICATORS IN THE WORLD AND IN THE RUSSIAN FEDERATION <i>A. V. Borisov, A. V. Bosov, and D. V. Zhukov</i>	89