

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

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

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

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

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

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

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

академик РАН И. А. Соколов — главный редактор

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

д.т.н. В. Н. Захаров

д.ф.-м.н. В. И. Синицын

проф., д.ф.-м.н. А. И. Зейфман

проф., д.т.н. И. Н. Синицын

проф., д.т.н. В. Д. Ильин

проф., д.ф.-м.н. В. Г. Ушаков

проф., д.т.н. К. К. Колин

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

проф., д.ф.-м.н. В. Ю. Королев

к.ф.-м.н. С. А. Христочевский

д.ф.-м.н. Р. В. Разумчик

Редакция

к.ф.-м.н. Е. Н. Арутюнов

С. Н. Стригина

© Федеральный исследовательский центр «Информатика
и управление» Российской академии наук, 2025

Журнал включен в базу данных Russian Science Citation Index (RSCI),
интегрированную с Web of Science, и в «Белый список» научных журналов

Журнал входит в систему Российского индекса научного цитирования (РИНЦ)

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

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

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

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

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

СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ

Том 35 № 1 Год 2025

СОДЕРЖАНИЕ

О роли академика И. А. Мизина в развитии информационных технологий и их применении в государственных интересах: к 90-летию со дня рождения И. А. Соколов, А. А. Зацаринный, В. Н. Захаров, К. К. Колин	3
Методы нормальной субоптимальной фильтрации в наблюдаемых неявных гауссовских стохастических системах И. Н. Синицын	41
Анализ сбоев по косвенным признакам А. А. Грушо, Н. А. Грушо, М. И. Забежайло, Д. В. Смирнов, Е. Е. Тимонина, С. Я. Шоргин	59
Исследование влияния механизма Лапласа на статистические характеристики исходных данных В. Н. Гридин, Б. Р. Салем, Д. С. Смирнов, В. И. Солодовников	71
Перспективные методы реализации инкрементального обновления материализованных представлений в современных реляционных системах управления базами данных С. А. Ступников, Н. А. Скворцов, Д. О. Брюхов	95
Интеграция электронного словаря с текстами параллельного корпуса: новый теоретический подход Д. О. Добровольский, И. М. Зацман	111
Гибридная интеллектуальная многоагентная система как модель малого коллектива специалистов для решения практических проблем И. А. Кириков, С. В. Листопад, С. Б. Румовская	125
Свойства и оптимизация самосинхронных схем В. Н. Захаров, Ю. А. Степченков, Ю. Г. Дьяченко, Н. В. Морозов, Л. П. Плеханов, Д. Ю. Степченков	149
Процессоры данных как новый компонент инфраструктуры дата-центров В. Б. Егоров	170
Об авторах	181
Правила подготовки рукописей статей	184
Requirements for manuscripts	188

О РОЛИ АКАДЕМИКА И. А. МИЗИНА В РАЗВИТИИ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ И ИХ ПРИМЕНЕНИИ В ГОСУДАРСТВЕННЫХ ИНТЕРЕСАХ: К 90-ЛЕТИЮ СО ДНЯ РОЖДЕНИЯ

И. А. Соколов¹, А. А. Зацаринный², В. Н. Захаров³, К. К. Колин⁴

Аннотация: Статья посвящена 90-летию юбилею академика И. А. Мизина, выдающегося ученого в области телекоммуникаций, информатики и информационных технологий. Рассмотрены основные научные и практические результаты, полученные им на протяжении 40 лет активной творческой деятельности. Игорь Александрович Мизин как главный конструктор обеспечил выполнение коллективом НИИ автоматической аппаратуры комплекса работ по разработке и внедрению первой отечественной территориально распределенной системы обмена данными (СОД) с пакетной коммутацией в интересах автоматизированной системы управления (АСУ) стратегического назначения. Под его руководством научные коллективы Института проблем информатики (ИПИ) РАН обосновали выбор информационных технологий для создания и развития информационных сетей в интересах органов государственной власти с учетом требований по защите информации. Показан вклад И. А. Мизина в становление и развитие российского информационного общества. Он вел активную научно-организационную работу, был членом бюро Отделения информатики, вычислительной техники и автоматизации (ОИВТА) РАН. Показано, что И. А. Мизин выступил создателем и признанным лидером отечественной научной школы в области теории и практики информационно-вычислительных и телекоммуникационных систем.

Ключевые слова: главный конструктор; академик; информационно-телекоммуникационные сети; информационные технологии; система обмена данными; центры коммутации сообщений; коммутация кодограмм; протоколы передачи информации; защита информации; каналы связи; органы государственной власти

DOI: 10.14357/08696527250101

EDN: FWNKEO

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

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

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

⁴Федеральный исследовательский центр «Информатика и управление» Российской академии наук, kolinkk@mail.ru



Игорь Александрович Мизин (1935–1999)

1 Введение

Двенадцатого апреля 2025 г. исполняется 90 лет со дня рождения академика РАН И. А. Мизина. Игорь Александрович Мизин был крупным ученым, обогатившим науку трудами первостепенного научного значения в области компьютерных наук, включая информационные технологии, вычислительную технику, средства передачи и защиты информации [1]. Он является создателем и признанным лидером отечественной научной школы в области теории и практики информационно-вычислительных и телекоммуникационных систем [2–6].

В последние 10 лет жизни (с 3 января 1989 г. по 15 июня 1999 г.) И. А. Мизин занимал должность директора Института проблем информатики АН СССР (с 1992 — Институт проблем информатики РАН, ИПИ РАН). В эти годы был главным редактором сборника «Системы и средства информатики», который в 2011 г. преобразован в одноименный журнал.

Игорь Александрович Мизин родился 12 апреля 1935 г. в г. Москве. Его отец, Александр Михайлович Мизин (1904–1968), уроженец деревни Солонино Ярославской области, связал свою жизнь с армией, воевал с басмачами в Средней Азии. Служил в инженерных войсках, участвовал в Великой Отечественной войне, закончил ее в Берлине, уволился в звании инженер-полковника. Мать Игоря Александровича, Александра Ануфриевна Мизина (Коваль) (1909–1979), родилась в селе Забужье Винницкой области. Высокопрофессиональная машинистка, Александра Ануфриевна работала в различных государственных учреждениях. К тому же она имела музыкальное образование, играла на фортепиано. Война застала семью Мизиных на Украине, в городе Полтаве. Игорь с мамой были эвакуированы на Урал в деревню Антипино. После эвакуации семья некоторое время жила в Ростове Великом, затем на Украине, в городе Ровеньки. В конце 1940-х гг. отца Игоря Александровича перевели на службу в Москву. Мизины

поселились в районе Абельмановской заставы, затем получили комнату на шоссе Энтузиастов. Учась в школе, Игорь Александрович активно занимался спортом — он был одним из сильнейших десятиборцев Москвы в своей возрастной категории. В то же время он отлично учился и в 1952 г. окончил среднюю школу № 464 Ждановского района г. Москвы с золотой медалью.

После окончания средней школы в 1952 г. был принят без экзаменов как золотой медалист в Военно-воздушную инженерную академию им. Н. Е. Жуковского (ВВИА), которую успешно окончил в феврале 1959 г. по специальности «Эксплуатация радиотехнических средств» и получил квалификацию «инженер по радиотехнике ВВС». В ВВИА им была выполнена и в 1959 г. опубликована первая научная работа, посвященная принципам построения систем однополосной радиосвязи [3].

В марте 1959 г. инженер-лейтенант И. А. Мизин был направлен в известную научно-промышленную организацию военно-промышленного комплекса — НИИ 101 Государственного комитета по радиоэлектронике (с 1964 г. — НИИ автоматической аппаратуры, в настоящее время НИИАА им. В. С. Семенихина). Здесь в период с 1959 по 1989 гг. он последовательно занимал должности инженера-лаборанта, инженера-исследователя, старшего инженера-лаборанта, руководителя группы, начальника лаборатории, начальника сектора, начальника отдела, начальника отделения, начальника комплексного научного отделения, начальника специального конструкторского бюро — заместителя директора по научной работе, начальника научно-тематического центра — заместителя директора по научной работе НИИАА (1979–1989). За эти годы он вырос в воинских званиях от лейтенанта-инженера до генерал-майора-инженера (1979 г.) [2].

В 1966 г. защитил диссертацию на соискание ученой степени кандидата технических наук на тему «Разработка метода обмена информацией по закрытым телеграфным каналам связи в командной системе управления специального назначения», а в 1972 г. — докторскую диссертацию на тему «Вопросы исследования и разработки информационных сетей территориальных автоматизированных систем управления». В 1975 г. И. А. Мизину присвоено ученое звание профессора по кафедре «Автоматизированные системы управления».

В декабре 1984 г. состоялись первые выборы в созданное в том году ОИВТА АН СССР. На них И. А. Мизин был избран членом-корреспондентом АН СССР по специальности «вычислительная техника». Начинается все более активное взаимодействие с учеными из академических кругов. В 1986 г. И. А. Мизин включен в состав Специализированного совета Д 003.56.01, созданного при ИПИ АН СССР под председательством директора института академика Б. Н. Наумова.

В июне 1988 г. после неожиданной смерти директора ИПИ Б. Н. Наумова институт остается без директора. Было ясно, что институт, завоевавший за четыре с половиной года значительный авторитет в стране и за рубежом, должна возглавить сильная фигура. В это время в результате происходящей в стране перестройки значительно усложняется деятельность научных организа-

ций оборонно-промышленного комплекса, в том числе и НИИАА. Происходит сокращение тематики, финансирования, уходят многие сотрудники. Со стороны руководства ОИВТА Игорю Александровичу было предложено подумать о возможности возглавить ИПИ АН. Ему предстояло принять очень непростое решение, ведь он всю жизнь проработал в «ящике», а тут надо было менять все окружение. После некоторых раздумий Игорь Александрович принял решение выдвинуть свою кандидатуру на уже объявленную вакансию директора ИПИ АН. Оформили это как выдвижение от базовой кафедры НИИАА в Московском институте радиотехники, электроники и автоматики (МИРЭА), которую он возглавлял. После этого он знакомится с подразделениями ИПИ АН, разбросанными в то время по многим точкам. На общем собрании ИПИ АН на голосование был представлен только один кандидат — И. А. Мизин, и его кандидатуру коллектив единодушно поддержал. Двадцатого декабря 1988 г. за него проголосовало Общее собрание ОИВТА, а 3 января 1989 г. Президиум АН СССР утвердил И. А. Мизина в качестве директора ИПИ АН. С 20 января 1989 г. И. А. Мизин приступил к работе в качестве директора ИПИ АН. В то же время он не прерывал активную работу в НИИАА, где до мая 1990 г. продолжал занимать должность заместителя директора по научной работе. Двадцатого апреля 1992 г. приказом Главнокомандующего Объединенными Вооруженными Силами СНГ маршала авиации Е. И. Шапошникова в звании генерал-майора был уволен в запас. Однако продолжал активное взаимодействие с органами военного управления по вопросам развития АСУ Вооруженных Сил (ВС).

На период директорства И. А. Мизина выпали, пожалуй, самые сложные годы в истории института. К моменту его прихода в институт тематика в значительной степени строилась с ориентацией на выполнение комплексных научных проектов по реализации Концепции создания новых поколений вычислительных систем в кооперации с научными организациями социалистических стран. Кроме того, ИПИ АН с 1985 г. был головной организацией Межотраслевого научно-технического комплекса (МНТК) «Персональные ЭВМ», в который входили ведущие в этом направлении научно-производственные организации СССР. В декабре 1989 г. И. А. Мизин был назначен генеральным директором МНТК «Персональные ЭВМ». Но с начала 1990-х гг. и особенно после распада СССР в 1991 г. все кооперации и связи начали стремительно распадаться. Внутри страны происходила переориентация с производства своих средств и систем вычислительной техники на покупные с Запада, приостанавливались запланированные стройки, резко сокращались научно-производственные организации, фактически ликвидировалась отраслевая наука. Как-то еще продолжали существовать академические институты. С распадом СССР ликвидировался и социалистический лагерь в Европе. Фактически прекратились налаженные научно-технические международные связи в рамках стран — членов СЭВ, который «приказал долго жить». В конце 1991 г. Академия наук СССР была преобразована в Российскую академию наук (РАН) и институт стал называться ИПИ РАН.

После прихода в ИПИ АН Игорь Александрович начал привносить определенные изменения в направления деятельности института, усилилась тематика, связанная с созданием и использованием систем и комплексов передачи данных в интересах различных заказчиков. Из состава института сразу выпал крупный филиал в Бердянске (Украина), на который возлагались достаточно серьезные научно-организационные задачи, и эти возникшие проблемы пришлось решать Игорю Александровичу. Позиция директора академического института позволяла уделять больше внимания проблемам развития новых информационных технологий и вопросам их использования в информационно-управляющих системах органов государственного управления и в системах специального назначения. В 1994 г. И. А. Мизин был назначен Генеральным конструктором системы информационного обеспечения управления государством «Юпитер», а в 1997 г. на него были возложены функции Генерального конструктора АСУ Вооруженными Силами Российской Федерации и одновременно Генерального конструктора НИИАА. В мае 1997 г. И. А. Мизин был избран действительным членом РАН (академиком) по ОИВТА по специальности «Вычислительная техника и элементная база».

Работая в ИПИ РАН, И. А. Мизин установил и поддерживал активные контакты с зарубежными коллегами из стран, с которыми ранее в институте не было взаимодействия. Он получил международную известность и снискал глубокое уважение многих ученых, профессоров и представителей деловых кругов в США, Германии, Италии, в странах Юго-Восточной Азии.

Игорь Александрович Мизин умело сочетал работу конструктора и руководителя крупномасштабных проектов с глубокими теоретическими исследованиями. Он автор более 190 научных работ в открытой печати, в том числе 12 монографий, 12 изобретений, а также огромного числа отчетов и закрытых проектов по созданию систем и комплексов различного назначения. Научные работы И. А. Мизина отличались глубиной и оригинальностью решений, сочетанием серьезного математического аппарата с инженерной интуицией и практической направленностью.

Академиком И. А. Мизиным сделан огромный научный и практический вклад в развитие информационных систем в рамках трех крупных направлений:

- (1) СОД с пакетной коммутацией в интересах АСУ стратегического назначения;
- (2) информационных технологий в интересах органов государственной власти;
- (3) технологий для информатизации общества.

2 Разработка и внедрение системы обмена данными с пакетной коммутацией в интересах автоматизированной системы управления стратегического назначения

Становление И. А. Мизина и как инженера-конструктора, и как ученого проходило в стенах НИИ автоматической аппаратуры, ведущего в стране

предприятия в области АСУ. В первые годы молодой военный инженер занимался исследованиями и разработкой методов повышения достоверности передачи информации по каналам связи различной физической природы, методов помехоустойчивого кодирования, вероятностных моделей дискретных каналов связи, которые были реализованы в нескольких комплексах повышения достоверности для автоматизированных систем военного назначения, в том числе в интересах противовоздушной обороны (ПВО), военно-воздушных сил (ВВС) и ракетных войск стратегического назначения (РВСН) [2]. Так, как специалист по авиационной технике с военно-инженерным образованием в 1960–1961 гг. он принял активное участие в государственных испытаниях комплексной АСУ наведением на воздушные цели сверхзвуковых самолетов-перехватчиков для тактического звена войск ПВО (система «Воздух-1»). Главный конструктор этой системы В. Я. Кравец и его заместитель В. В. Конашев поручили И. А. Мизину разработку технических решений по передаче командной информации на борт самолета-перехватчика по радиоканалам. Он успешно справился с поставленными задачами, и в 1962 г. В. Я. Кравец, уже как Главный конструктор АСУ в интересах нового вида Вооруженных Сил — РВСН, поручил ему разработку средств передачи данных и повышения достоверности для передачи командной информации по каналам связи низкого качества. Во вновь созданном отделе АСБУ РВСН начальник отдела В. В. Конашев (будущий директор НИИАА) доверил И. А. Мизину должность начальника лаборатории аппаратуры повышения достоверности передачи данных. В 1962 г. коллективом лаборатории под руководством И. А. Мизина были разработаны средства повышения достоверности и передачи данных для АСУ РВСН [7].

Целеустремленность и работоспособность молодого военного инженера, а также достигнутые им научно-практические результаты были замечены руководством института, и в 1964 г. Игоря Александровича назначают начальником 11 отдела (аппаратуры передачи данных). Игорь Александрович Мизин собирает в отделе своих ближайших соратников: Г. К. Храмешина, Л. С. Уринсона, С. К. Муравьева, В. В. Муравьева, Ю. Я. Кислова, Г. С. Вильшанского, Л. А. Полянского, Б. С. Дудкина, Ю. А. Овчинникова, Л. И. Мавродиadi, Г. И. Батурина и др. Забегая вперед, отметим, что все перечисленные сотрудники в последующем были выдвинуты И. А. Мизиным на руководящие должности (сектор, отдел, отделение) в составе Научно-тематического центра систем передачи данных (НТЦ СПД) и много лет возглавляли важнейшие разработки различных комплексов и средств [2, 6].

Вот как это вспоминает Л. Б. Шукин, ветеран НИИАА, который хорошо знал И. А. Мизина с первых дней работы в НИИАА: *«Всего за два года им был создан отдел специалистов, которые связали свою творческую деятельность под его руководством с созданием теории и практики систем и комплексов передачи данных (с повышением достоверности) по каналам связи с ограниченной пропускной способностью и высоким уровнем помех. Непосредственная работа с ним всегда оказывалась деловой и конструктивной, представляющей всем*

разработчикам инициативу в решении конкретных технических вопросов после принципиальных договоренностей с ним как с главным конструктором средств передачи данных в автоматизированной системе 15Э1» [8].

Вспоминает В. И. Петров, выпускник Московского авиационного института (МАИ) 1964 г., отобранный И. А. Мизиным в 11 отдел: *«И. А. Мизин прекрасно разбирался в людях, говорил кратко, понятно и убедительно. После защиты диплома команда МАИ (включая меня), обосновалась на Кировской. Гигантские задачи решал 11 отдел. Замечательные талантливые молодые ребята, стариков не было! Игорь Александрович сформировал уникальный коллектив! Мизин был Человек, он переживал и болел за своих сотрудников, был к нам близок, играл с нами в шахматы, волейбол, ездил на овощные базы, обладал чувством юмора. Отлично руководил, всегда выслушивал наши точки зрения, прежде чем принять решение, проводил технические совещания» [9].*

Судьбоносным для И. А. Мизина стал 1967-й год, когда Постановлением ЦК КПСС и Совета Министров была задана разработка АСУ стратегического звена, а НИИАА этим Постановлением определен головной организацией. При формировании коллектива исполнителей этой важнейшей работы ему была поручена разработка СОД, призванной обеспечить информационный обмен данными с требуемыми характеристиками между объектами, территориально распределенными по территории всей страны. Игорь Александрович Мизин был назначен Главным конструктором СОД [2].

Особенность разработки СОД заключалась в необходимости обеспечения беспрецедентно высоких вероятностно-временных характеристик доведения информации между объектами, территориально распределенными по территории всей страны, в интересах обеспечения эффективного функционирования командной системы управления. Однако результаты проведенных предпроектных обследований объектов высшего звена управления показали, что при построении системы по классическим принципам иерархической структуры комплексы средств автоматизации (КСА) на объектах высшего звена управления превращались в огромные многоканальные узлы связи, обеспечить обслуживание и эксплуатацию которых не представлялось возможным. Требовался принципиально другой подход к построению СОД.

И тогда в качестве выхода из создавшегося положения И. А. Мизиным совместно с Л. С. Уринсоном и Г. К. Храмешиным был предложен новый для отечественной практики метод реализации информационного обмена между объектами АСУ, состоящий в использовании сетевой структуры с автоматическими центрами коммутации сообщений (ЦКС) [10]. Это предложение было основано на анализе опыта создания сети ARPANET [11–13].

Такое решение кардинально разрешало «объектовые проблемы» и придавало системе функциональную гибкость, но вместе с тем требовало соответствующего вычислительного ресурса и многофункционального программного обеспечения (ПО) для создания такого сложного комплекса, как ЦКС. Это была своеобразная «плата» за переход к созданию СОД на принципах коммутации сообщений.

Для успешного решения этих проблем сравнительно немногочисленный коллектив отдела И. А. Мизина, который во второй половине 1960-х гг. начинал разработку системотехнических решений для создания СОД, постоянно увеличивался и к началу государственных испытаний (1978 г.) превратился в мощнейший коллектив высококвалифицированных опытных ученых, конструкторов, инженеров и специалистов — НТЦ СПД.

Игорь Александрович Мизин благодаря своим незаурядным организаторским способностям и человеческим качествам сумел подобрать и подготовить достойных руководителей всех основных направлений разработки [2, 6].

Отметим несколько «ключевых» фигур в команде создателей СОД.

Уринсон Леонид Савельевич, кандидат технических наук, возглавлял системное отделение, лично руководил разработкой комплексов и средств передачи данных. Умер, к сожалению, в 1977 г. совсем молодым в возрасте 40 лет, в полном расцвете сил.

Храмешин Геннадий Кузьмич, доктор технических наук, полковник (выпускник Ленинградской военно-воздушной инженерной академии им. А. Ф. Можайского), заместитель Главного конструктора, возглавлял разработку системотехнических вопросов построения СОД и комплексов средств передачи данных. Заведующий базовой кафедрой МИРЭА. Соавтор известной монографии [10]. Он успешно заменил Л. С. Уринсона на должности начальника системного отделения.

Матюхин Николай Яковлевич, выпускник МЭИ, член-корреспондент АН СССР (1979), один из создателей первых отечественных ЭВМ, ученик выдающегося советского ученого в области вычислительной техники члена-корреспондента АН СССР И. С. Брука. Разработчик первых отечественных машин для научных исследований М-1, М-2, М-3 [14]. В НИИАА с 1957 г. [2]. В команду И. А. Мизина Н. Я. Матюхин вошел в начале 1970-х гг. уже с огромным научным и практическим опытом. Главный конструктор ЦКС, начальник крупного отделения.

Тамошинский Анатолий Васильевич, доктор технических наук, заменил Н. Я. Матюхина на посту Главного конструктора ЦКС после его преждевременной смерти (в 1984 г.), а в последующем стал Главным инженером НИИАА и затем ученым секретарем этого института.

Березин Владимир Николаевич, доктор технических наук, на протяжении двух десятилетий (1970–1980-е гг.) был заместителем Главного конструктора и руководителем крупного подразделения программистов в составе НТЦ СПД. Особо отметим, что В. Н. Березин стал преемником И. А. Мизина после его ухода в Институт проблем информатики РАН и в самые тяжелые 1990-е гг. в течение более десяти лет достойно выполнял обязанности руководителя НТЦ СПД и Главного конструктора СОД [6].

Кулешов Александр Петрович, выпускник мехмата МГУ, один из ведущих разработчиков уникального протокола сетевого уровня, начальник сектора в коллективе В. Н. Березина. В 1980-е гг. стал главным инженером, затем первым

заместителем начальника НТИЦ СПД и, по существу, в эти годы одним из ближайших помощников И. А. Мизина. В настоящее время академик РАН, с 2016 г. ректор Сколковского института науки и технологий.

Муравьев Сергей Константинович, кандидат технических наук, капитан первого ранга, руководил отделением, в котором разрабатывались комплексы оконечных средств СОД (КОССОД) для подключения к ЦКС комплексов средств автоматизации командной системы боевого управления (КСБУ) и других автоматизированных подсистем. Отделением был разработан ряд унифицированных КОССОД для объектов различных видов базирования (стационарных, подвижных, воздушных).

Чипашвили Шота Шотович, один из ближайших соратников И. А. Мизина, кандидат технических наук, полковник, успешно решал многочисленные организационно-технические вопросы по созданию и поддержанию стенда Главного конструктора, а также по заблаговременной подготовке объектов заказчика для установки на них ЦКС и обеспечению условий для их эксплуатации.

В 1979 г. команду И. А. Мизина существенно усилил перешедший в НИИАА из 16 ЦНИИ МО доктор технических наук Богатырев Вячеслав Андреевич, один из ведущих системотехников в области систем связи с коммутацией пакетов [6, 15].

В результате коллективом НТИЦ СПД под руководством И. А. Мизина был успешно решен целый ряд новых научно-теоретических и системотехнических задач, обеспечивших выполнение множества требований со стороны АСУ к информационному обмену. Основные из них [2, 6]:

- обоснование структуры циркулирующих в СОД сообщений, включая формализацию служебной части, с учетом требований объектов управления;
- обоснование системы адресования сообщений, передаваемых через СОД, с обеспечением вариантов их доведения (циркулярно, циркулярно-избирательно, избирательно, по спискам, многоадресные);
- оптимальное кодирование передаваемой в сообщениях формализованной служебной и содержательной информации;
- создание совокупности (стека) алгоритмов, протоколов и интерфейсов, обеспечивающих передачу через СОД всех необходимых видов информации с заданными вероятностно-временными характеристиками;
- обеспечение защиты информации, циркулирующей в СОД;
- реализацию задач маршрутизации сообщений на вычислительных комплексах центров коммутации;
- обеспечение высоких показателей достоверности передачи сообщений по каналам связи, в том числе низкого качества;
- обеспечение высоких показателей надежности трактов передачи сообщений в СОД между объектами управления;

- создание в составе СОД подсистемы контроля функционирования и управления ее работой;
- создание специализированных аппаратно-программных комплексов для центров коммутации пакетов (ЦКП);
- создание ряда аппаратно-программных КОССОД, устанавливаемых на объектах управления АСУ — абонентах СОД;
- создание комплекса ПО для ЦКП.

Успешная реализация всех перечисленных выше задач позволила своевременно решить весь комплекс проблем по вводу СОД в эксплуатацию, которые включали:

- проведение различных видов испытаний программно-аппаратных средств, комплексов и опытного участка СОД в целом;
- отладку средств и комплексов СОД непосредственно на объектах эксплуатации;
- ввод СОД в штатный режим дежурства, т. е. круглосуточного обслуживания системы управления.

При создании СОД НИИАА выполнял основной объем работ, однако надо отметить большой вклад и кооперации предприятий-соисполнителей. Среди них: Ульяновский НПО «Марс» (главный конструктор — В. В. Алексейчик, затем — В. И. Кидалов), НПО «Красная Заря» (главный конструктор — М. И. Ветчинкин, затем — В. И. Мирошников), Ереванский НИИ математических машин (главный конструктор — Р. В. Атоян), КБ завода «Сигнал» (г. Кишинев, главный конструктор — И. Б. Петяшин), НИИССУ (Научно-исследовательский институт систем связи и управления, директор — В. А. Соколов), МНИИРС (Московский научно-исследовательский институт радиосвязи, директор — А. П. Биленко), НИИ автоматики (Москва, директор — А. Ф. Алексеев), Пензенский научно-исследовательский электротехнический институт (ПНИЭИ, директор — А. В. Старовойтов), Горьковский НИИ радиосвязи (директор — Е. Л. Белоусов), Электромеханический завод (г. Свердловск), Тамбовский завод «Тамбоваппарат», Рязанский завод «Красное знамя», завод САМ (Москва), монтажные предприятия «Каскад» и др. [2].

Отметим, что И. А. Мизин в своей работе всегда получал помощь и постоянную деловую поддержку со стороны руководства НИИАА в лице В. С. Семенихина и В. В. Конашева [2].

Особо необходимо подчеркнуть поддержание И. А. Мизиным на протяжении многих лет конструктивных деловых отношений с Заказчиком КСБУ и СОД в лице Управления начальника связи Вооруженных Сил [16], а также ведущими военными научно-исследовательскими институтами, прежде всего с 16 ЦНИИИ связи Минобороны [17, 18].

Создание первой в СССР телекоммуникационной системы с коммутацией пакетов для командной АСУ и ее постановка в режим круглосуточной эксплуатации

стали крупным научно-техническим достижением отечественной системотехники. Игорь Александрович Мизин как главный конструктор в 1981 г. удостоен Ленинской премии, а ряд его сотрудников — Государственной премии, награждены орденами и медалями.

Новые решения, заложенные в основу построения СОД, оказались эффективными и перспективными, они обеспечили в процессе эксплуатации возможность ее эволюционного развития и превращения в базовую СОД (БСОД) как основу объединенной СОД (ОСОД).

Именно поэтому БСОД, созданная коллективом И. А. Мизина, поставленная в 1980 г. на опытную эксплуатацию, а с 1985 г. принятая на боевое дежурство, успешно функционирует более 40 лет в интересах нескольких десятков автоматизированных систем в стратегическом звене управления.

Более того, по истечении нескольких десятилетий выяснилось, что многие системотехнические подходы, реализованные И. А. Мизиным и его командой в ходе создания СОД, предвосхитили ряд идей, которые легли в основу передовых зарубежных разработок в этой области (однако в силу закрытого характера работ НИИАА эти положения остались неизвестными международной научной общественности).

Во многом это объясняется высоким уровнем научной квалификации коллектива разработчиков СОД. В руководимом И. А. Мизиным НТЦ СПД не ограничивались только разработками и внедрением комплексов обмена данными на объектах заказчика, много внимания уделялось развитию теоретических вопросов построения больших информационно-телекоммуникационных систем (ИТКС). Этому способствовали работы по математическому и имитационному моделированию разрабатываемых системных, алгоритмических и технических решений, а также экспериментальных исследований с использованием стенда главного конструктора. И здесь надо отметить роль лидера коллектива: Игорь Александрович как Главный конструктор в условиях колоссальной загруженности повседневными проблемами разработки СОД нашел время для научных исследований и сумел сравнительно рано защитить кандидатскую (в 1966 г.), а затем и докторскую (в 1972 г.) диссертацию. Вспоминает профессор И. А. Ушаков (член первого диссертационного совета НИИАА, который создал В. С. Семенихин, сотрудник НИИАА, затем ВЦ АН СССР): *«Игорь защищал кандидатскую диссертацию в 1966 г. на нашем Совете. Я помню его блестящую защиту. Игорь представил отличный технический проект, в котором все — от теоретического обоснования, инженерных расчетов, конструктивных решений вплоть до макетного образца — было сделано им самим»* [19].

Личным примером И. А. Мизин высоко поднял планку квалификационных требований к руководителям структурных подразделений НТЦ СПД. Практически все они уже к середине 1980-х гг. стали дипломированными специалистами: доктора технических наук Г. К. Храмешин, В. Н. Березин, А. П. Кулешов, А. В. Тамошинский, кандидаты технических наук Л. С. Уринсон, С. К. Муравьев, Л. А. Полянский, Л. И. Мавродиadi, А. Б. Залкинд, Г. С. Вильшанский,

В. В. Муравьев, Ю. В. Михеев, В. Н. Веселов, Ю. Я. Кислов, Ю. А. Овчинников, Г. А. Кабатянский, С. А. Осмоловский, Ф. И. Фишбеин. В конце 1980-х гг. кандидатские диссертации успешно защитили В. Н. Баев, В. Т. Самохин, М. Б. Куперман, П. А. Платонов, Э. А. Верещагин и др.

Отдельно отметим, что главный конструктор ЦКС Н. Я. Матюхин в 1979 г. был избран членом-корреспондентом Академии наук СССР. В 1979 г. научный потенциал НТЦ СПД усилился В. А. Богатыревым, который после успешной защиты докторской диссертации перешел в НИИАА из 16 ЦНИИИ [15].

Важно, что все диссертации специалистов НТЦ СПД были основаны на результатах (технические решения, методы, алгоритмы), реально реализованных в комплексах СОД (устройства преобразования сигналов, устройства повышения достоверности, вычислительный комплекс коммутации, средства отображения и др.).

Еще одним направлением деятельности И. А. Мизина была преподавательская работа на базовой кафедре МИРЭА, которую он возглавлял. Вместе с ним высококвалифицированные инженерные кадры для НТЦ СПД на базовой кафедре готовили Г. К. Храмешин, В. Н. Березин, Н. Я. Матюхин, А. Б. Залкинд, Г. С. Вильшанский, В. В. Муравьев и др.

Отметим также, что по проблематике сетей передачи данных И. А. Мизиным в соавторстве с коллегами (Л. С. Уринсон, Г. К. Храмешин, В. А. Богатырев, А. П. Кулешов и др.) в период разработки СОД был опубликован ряд фундаментальных научных работ [10, 20–23], на которых учились несколько поколений исследователей и разработчиков в области ИТКС.

Система обмена данными стала образцом системного подхода к созданию крупных территориально распределенных телекоммуникационных систем [6, 17].

3 Создание информационных сетей в интересах органов государственной власти

С приходом И. А. Мизина в ИПИ РАН стало развиваться новое научное направление исследований, которое получило название «Информационно-телекоммуникационные системы и технологии», и была сформирована научная школа отечественных специалистов в области создания сложных ИТКС [22–24]. К этому времени он обладал колоссальным опытом разработки и внедрения крупных информационных систем. В нем органично сочетались качества ученого, способного воспринять и правильно оценить новое, и инженера-конструктора, нацеленного на достижение конкретных целей. Игорь Александрович придал этому направлению яркую практическую направленность. При этом сложившееся в начале 1990-х гг. в России положение в области информатики позволило И. А. Мизину в полной мере реализовать свой огромный научный потенциал и редчайший организаторский талант.

В эти годы в стране появились небывалые возможности по доступу и применению широчайшего спектра информационных технологий, разработанных ведущими бизнес-структурами США, Германии, Франции и других ведущих стран. В России наступил переломный момент в становлении и использовании информационных технологий в различных сферах: промышленности, науке, образовании, а также в государственном управлении.

В сентябре 1994 г. вышел указ Президента РФ, поставивший актуальнейшую задачу модернизации информационных и телекоммуникационных систем органов государственной власти на базе использования последних достижений в этой отрасли. Этим же указом И. А. Мизин был назначен Генеральным конструктором, а Институт проблем информатики РАН определен головной организацией. Предстояла не только сложная, но и во многом пионерская работа. Требовалось с новых позиций изучить процессы управления, вычленить основные функциональные задачи, сформулировать количественные и качественные характеристики, необходимый набор исходной информации и, наконец, создать комплексы аппаратно-программных средств (в том числе сетевого исполнения), позволяющие с требуемыми характеристиками автоматизировать эти процессы и создать компьютерные системы поддержки принятия решений и автоматизации процессов государственного управления.

Для успешного решения такой сложной практической задачи необходимо было обосновать теоретические подходы к описанию класса крупномасштабных информационных систем двойного применения как территориально структурированных мультисетевых систем национального масштаба, совмещающих функции специальных систем и систем общего пользования. При этом, с одной стороны, необходимо было внедрить самые современные и перспективные технологии и реализующие их средства, а с другой — обеспечить их преемственность с уже функционирующими системами. Это были очень сложные, масштабные и ответственные государственные задачи. И здесь И. А. Мизин в полной мере проявил свои лучшие качества: системность, широту взглядов, целеустремленность. Именно в процессе этой работы были сформулированы фундаментальные принципы автоматизации процессов государственного управления в применении к России, разработаны системотехнические решения построения автоматизированных информационных систем, включая конкретные аппаратно-программные решения. Разработанные принципы и подходы и сегодня остаются актуальной методологической основой для многих разработчиков в нашей стране.

Отметим, что И. А. Мизин обладал редчайшим талантом формировать высококвалифицированные коллективы для выполнения сложных и важных научно-практических задач. Он ярко это доказал в НИИИАА при создании СОД.

Не менее успешно свои организаторские способности И. А. Мизин проявил и в ИПИ РАН. Для решения задач по **модернизации информационных систем органов государственной власти** потребовалось сформировать новые научные коллективы. Их основу составили опытные сотрудники ИПИ РАН (К. К. Колин, Ю. В. Гайкович, И. Н. Сеницын, И. И. Быстров, Л. А. Калиниченко,

В. А. Козмидиади, С. М. Рождественский, В. Д. Ильин, И. М. Адамович, А. И. Мартыненко и др.), которые выполняли прежде всего исследовательские работы. Вместе с тем для решения большого объема задач, требующих выполнения работ по конструированию и разработке (рабочая и конструкторская документация, опытные образцы, все виды испытаний и т. п.), И. А. Мизин пригласил целый ряд сотрудников НИИАА с опытом разработки СОД (А. И. Темнов, В. Н. Веселов, Э. В. Киселев, М. Е. Медведев, М. А. Чудаков, В. И. Гребенщиков, А. А. Кураков, В. В. Васильев, И. В. Сергеев и др.). Наряду с этим к выполнению новых работ были привлечены совсем молодые сотрудники ИПИ РАН, выпускники инженерных вузов — МАИ, МИРЭА и др. Среди них отметим В. И. Синицына, Э. Р. Корепанова, В. А. Кондрашева, М. М. Гершковича, И. В. Макаренкову, Е. Н. Федотову. Все перечисленные сотрудники стали ведущими специалистами ФИЦ ИУ РАН по различным вопросам информационных систем и технологий.

Игорь Александрович Мизин большое внимание уделял научной обоснованности предлагаемых решений, в частности **методам математического моделирования ИТКС**. По его инициативе в 1992 г. в ИПИ РАН было создано новое подразделение — Отдел исследований и разработки принципов построения информационно-вычислительных систем, одной из задач которого стало математическое моделирование ИТКС. Основу отдела составили сотрудники, пришедшие в ИПИ РАН из расформированного совместного отдела НИИССУ и ВЦ АН СССР (С. Е. Шибанов, И. А. Соколов, С. Я. Шоргин, И. В. Лызлова, С. В. Антонов). В последующем вопросами математического моделирования ИТКС активно занимались А. В. Печинкин, В. В. Чаплыгин, М. Г. Коновалов, Я. М. Агаларов, Р. В. Разумчик.

Основным направлением исследований в данной области было решение научных задач по разработке методов и средств моделирования реальных телекоммуникационных систем. В частности, проводились фундаментальные теоретические исследования, разрабатывались математические и программные средства и модели, их экспериментальная апробация, анализ и оптимизация характеристик ИТКС. Среди наиболее значимых результатов следует отметить [24–26]:

- комплекс методов, алгоритмов и программ для расчета характеристик телекоммуникационных сетей;
- математическая и имитационная модели многоскачковых пакетных радиосетей;
- модели транспортного уровня для расчета и оптимизации характеристик доставки многопакетных сообщений;
- математические методы и алгоритмы для расчета общесетевых характеристик процессов передачи для АТМ-сети;
- модель транспортной сети, учитывающая зависимость среднего времени передачи сообщений от управляющих параметров транспортного протокола;

- алгоритм выбора параметров транспортного протокола, оптимизирующих среднее время передачи через транспортную сеть;
- концептуальная, математическая и программная модель телекоммуникационной сети, использующей технологию передачи Frame Relay (данный проект выполнялся в рамках сотрудничества научного коллектива с компанией SOTAS, США).

Остановимся на научно-практических результатах, полученных по некоторым направлениям.

Одним из дискутируемых вопросов был **выбор архитектуры КСА**, или, в терминологии, предложенной И. А. Мизиным, комплексов средств информатизации (КСИ) [22]. Рассматривались две основные альтернативы: централизованная архитектура, основанная на одном мощном центральном вычислителе с подключенными к нему терминалами, и децентрализованная клиент-серверная архитектура на базе интеллектуальных автоматизированных рабочих мест (АРМ) и целевых серверов, объединенных локальной вычислительной сетью (ЛВС). В результате углубленных аналитических исследований, включая анализ опыта крупнейших зарубежных компаний, а также практических экспериментов, была обоснована ориентация на децентрализованную клиент-серверную архитектуру. Отказ от широко распространенной в течение многих лет в Минобороны и других силовых ведомствах практики использования терминальных сетей на базе универсальных ЭВМ (ЕС ЭВМ, ИБМ) был обусловлен тем, что произошла коренная трансформация концептуальных взглядов на использование средств вычислительной техники и телекоммуникаций. Сущность этой трансформации состояла, во-первых, в доступности мощных персональных ЭВМ (ПЭВМ) для подавляющего большинства пользователей, во-вторых — в возможностях ПЭВМ решать широкий класс задач (а не только узкоспециализированных задач, решаемых высококвалифицированным персоналом). Поэтому при создании названных выше систем задача прежде всего состояла в обеспечении пользователя (специалиста) всем набором программных средств, позволяющих ему заниматься своей профессиональной деятельностью. Такой набор программных средств мог создаваться путем адаптации компонентов уже готового программного обеспечения независимых производителей под требования конкретного заказчика. При таком подходе сроки создания информационных систем существенно сокращались. Это, в свою очередь, означало, что перед создателями системы ставилась не только задача интеграции на одном АРМ системного ПО разных производителей для его совместной работы, но и задача поддержки последующего расширения перечня устанавливаемых программных средств. Имевшиеся на тот момент тенденции развития архитектур программных систем однозначно указывали на необходимость выбора архитектуры приложения, в которой прикладные и пользовательские сервисы реализованы на клиентской рабочей станции, а данные централизованно хранятся и обрабатываются на сервере.

Другим важным вопросом стал выбор технологии организации ЛВС. В отличие от сегодняшней ситуации, когда безраздельно «господствует» протокол Ethernet, в начале 1990-х гг. применялись и Ethernet (в двух модификациях — «тонкий» и «толстый»), и Token Ring, и FDDI, и ряд других протоколов. В результате анализа отечественного и зарубежного опыта, сравнительных характеристик по ряду показателей (пропускная способность, время доведения и др.), включая стендовое макетирование, а также массовости применения (этот фактор сыграл значительную роль) предпочтение было отдано технологии Ethernet, в дальнейшем с использованием коммутаторов ЛВС.

Вместе с тем было понятно, что эффективность решения поставленных задач во многом будет определяться не только архитектурой и технологией ЛВС, но также организацией хранения информации, методами ее поиска и извлечения и в конечном итоге **управлением базами данных (БД)**.

Жесткие ограничения по производительности процессоров начала 1990-х гг., ограниченные объемы оперативной памяти в условиях необходимости реализации сложных алгоритмов обработки данных часто приводили коллективы разработчиков к собственным решениям по способам структуризации данных в рамках сетевой модели, построению уникальных индексных систем, собственных языков манипулирования данными, их интерпретаторов и т. п. Особенно остро эти проблемы стояли при обработке неформализованной текстовой информации на русском языке.

В этих условиях в ИПИ РАН были обоснованы стратегические решения по унификации подходов к хранению данных в рамках реляционной модели и использованию языка SQL как унифицированного средства взаимодействия с **системами управления БД (СУБД)** (Л. А. Калиниченко, И. М. Адамович, Э. Р. Корепанов, О. И. Волков, М. М. Гершкович). В результате этой работы при наличии нескольких десятков вариантов в 1995 г. в качестве основных были выбраны три линейки СУБД:

- (1) СУБД компании Oracle (один из лидеров рынка на тот момент) для мощных серверов, обрабатывающих большие объемы фактографической и документальной информации с интенсивным потоком транзакций;
- (2) персональная СУБД Microsoft Access как альтернатива многочисленным линейкам СУБД форматов dBASE и Paradox для создания БД для небольших групп пользователей;
- (3) СУБД Microsoft SQL Server для технологических и учетных БД.

Впоследствии на основе этих СУБД было разработано большое число приложений, которые с небольшими изменениями работают и в настоящее время. Отметим, что и сейчас, спустя 30 лет, рекомендованные линейки СУБД остаются одними из востребованных для применения в своих сегментах. Так, согласно последнему рейтингу DB-Engines (декабрь 2024 г.), который ранжирует СУБД по популярности, первое место уверенно занимает СУБД Oracle, Microsoft SQL Server — на третьем, Microsoft Access — 12 из более чем 400 систем. Заметим,

что их многочисленные конкуренты в 1990-е гг. (Sybase SQL, Watcom SQL, Informix, Interbase, Paradox, FoxPro) практически исчезли с российского ИТ-рынка: Informix — 36 место, Interbase — 75; вне рейтинга — Sybase (поглощена SAP, затем упразднена в 2011), Paradox — последнее обновление в 2009 г., FoxPro — поддержка прекращена в 2007 г. [27].

Создаваемые крупномасштабные системы требовали новых, современных, решений по контролю их состояния, **управлению функционированием и эксплуатацией**. Здесь необходимо отметить, что особое внимание к решению этих вопросов было обусловлено опытом И. А. Мизина, приобретенным в период создания СОД. Разработанная в составе СОД система контроля состояния и управления играла значительную роль в обеспечении непрерывного функционирования СОД с заданными показателями.

В 1990-е гг. появился целый ряд новых продуктов этого класса — сетевых платформ управления (к тому времени уже были специфицированы и приняты в качестве международных стандартов протоколы SNMP). Наиболее известными были продукты фирмы Hewlett Packard (HP) — линейка продуктов HP Open View и Computer Associates (CA) — программный комплекс UniCenter.

В ИПИ РАН были приобретены оба продукта, построены тестовые и испытательные стенды, проведена большая исследовательская работа на различных конфигурациях серверного оборудования и средствах телекоммуникаций, операционных системах, СУБД, разнообразном общесистемном и прикладном ПО. Специалисты института прошли обучение на курсах фирм HP и CA.

В результате было принято решение: разрабатывать свою систему управления с включением в нее отдельных модулей из линейки HP Open View (в части управления аппаратными средствами и контроля трафика) и модулей из состава продукта UniCenter в части управления ПО (И. А. Соколов, И. В. Сергеев, Р. Н. Шагин, И. В. Боголюбов, С. В. Антонов, А. В. Босов, А. В. Борисов, Э. В. Киселев, В. И. Синицын).

Отдельно следует коснуться **проблемы сопряжения систем**, создаваемых на базе современных технологий, и систем, уже созданных и функционирующих на территории РФ. Здесь отметим две работы.

Первая из них — сопряжение с наиболее массовыми и территориально распространенными сетями того времени: телеграфными сетями АТ-50 и Телекс. Такое решение позволяло повысить вероятность доведения сообщений ограниченного объема (команды, сигналы, распоряжения и т. п.) с учетом увеличения множества возможных путей передачи информации за счет разветвленности телеграфной сети. С учетом таких потенциальных возможностей этих сетей в ИПИ РАН был разработан специальный комплекс сопряжения с сетями абонентского телеграфирования АТ-50 и Телекс для обеспечения обмена телеграфными сообщениями между абонентами объектовых ЛВС (Ethernet) и абонентами телеграфных сетей АТ-50 и Телекс (В. И. Гребенщиков, Р. Б. Чавтараев).

Вторая работа — сопряжение с объектами управления, подключенными к БСОД Вооруженных Сил РФ, которая к началу 1990-х гг. служила основной

транспортной сетью для АСУ ВС РФ. Сопряжение обеспечивало возможность информационного обмена с объектами военных систем. Для реализации такого сопряжения потребовалась разработка специального комплекса и правил по преобразованию принятых в БСОД форматов кодограмм в форматы протокола Ethernet, что позволило обеспечить информационное взаимодействие абонентов сети Ethernet модернизируемой системы органов государственной власти и абонентов КСА автоматизированных систем военного назначения. И такой, весьма непростой, комплекс сопряжения был создан.

В процессе проектирования были рассмотрены несколько вариантов реализации этого сопряжения. В качестве основного был выбран вариант, при котором комплекс сопряжения модернизируемой системы органов государственной власти подключался по выделенным каналам связи к центрам коммутации БСОД в виде виртуального объекта управления, обладающего необходимыми правами и характеристиками.

Основная сложность реализации такого вида сопряжения заключалась в том, что СОД была реализована исключительно на специализированных программно-аппаратных средствах и опыт сопряжения с ней систем, построенных на базе ЛВС, полностью отсутствовал. Для решения этой сложной технической задачи были использованы несколько двухканальных интеллектуальных контроллеров и специально разработанное ПО. Несмотря на отсутствие опыта выполнения подобных работ, средства сопряжения были успешно реализованы с использованием компонентов, развернутых на стендах главных конструкторов сопрягаемых систем. В процессе испытаний были проверены возможности обмена формализованной и неформализованной информацией между абонентами сопрягаемых систем с заданными характеристиками.

Очевидно, что успешному решению этой задачи способствовало то, что И. А. Мизин как Главный конструктор СОД прекрасно владел существом и особенностями технических решений, реализованных в БСОД АСУ ВС РФ.

Совершенно новые задачи приходилось решать в области **информационной безопасности**. Одновременно с формированием нормативной базы (совместно с Федеральным агентством правительственной связи и информации при Президенте Российской Федерации (ФАПСИ), Федеральной службой безопасности (ФСБ) и Министерством обороны) были испробованы различные подходы и решения, в том числе и не нашедшие дальнейшего применения. В конце концов вышли на использование решений, которые определяют современное состояние защищенности государственной тайны в автоматизированных системах. Разработанные средства позволили обеспечить защиту информации при работе пользователей в доменной структуре, обмене сообщениями электронной почты, доступе к информации в БД и при совместной работе над документами. Средства криптографической защиты реализуют шифрование и имитозащиту файлов, функции электронного замка, шифрование внутрисетевого трафика и защиту внешних каналов связи. Впервые в изделиях были реализованы криптоалгоритмы семейства «Метель», разработаны специальные считывающие устройства, в качестве но-



На полевых учениях войск ФАПСИ в период 1996–1997 гг. Слева направо: первый заместитель Генерального директора ФАПСИ генерал-полковник Н. Н. Андреев; Генеральный директор ФАПСИ генерал-полковник А. В. Старовойтов; генерал-майор И. А. Мизин; секретарь Совета обороны Российской Федерации Ю. М. Батурин

сителей ключевой информации применена Российская интеллектуальная карта, реализованы различные типы интерфейсов, включая высокоскоростные оптоволоконные. Разработанные изделия успешно прошли сертификацию ФСБ и широко применяются в защищенных сетях органов государственной власти РФ.

Оперативность и правильность принятия решений руководством страны во многом зависит от наглядности и качества представления данных, что требовало широкого внедрения технологий **геоинформационных систем (ГИС)**, позволяющих эффективно связывать картографическую основу и фактографические БД, а также быстро отображать огромный объем накопленных данных в пригодном для восприятия виде (А. И. Темнов, А. А. Кураков, М. А. Чудаков, Л. С. Берштейн, И. Н. Розенберг). Эффективность ГИС в процессе оперативного принятия решений была настолько очевидной, что под них создавались специализированные средства отображения информации — «видеостены», которые позволяли получать приемлемые характеристики по разрешению, яркости и размеру картографического изображения при достаточно скромных характеристиках единичного технического элемента (проектора, видеокуба). В состав интегрированного программно-аппаратного комплекса, реализующего функции представления информации для принятия решения, входил экран коллективного пользования проекционного типа на базе видеокубов, индивидуальные средства отображения и аппаратно-программные средства поддержки.

Большое внимание было уделено исследованиям технологий пакетной коммутации на базе протоколов серии X.25 (X.3, X.28, X.29, X.224 и т.д.), определяемых рекомендациями Международного консультативного комитета по телефонии и телеграфии. Эти технологии получили широкое распространение

в 1990-е гг. в связи с развитием процессов создания телекоммуникационных сетей в интересах различных структур, которые разрабатывались на оборудовании зарубежного производства. И здесь важно отметить усилия И. А. Мизина по практическому применению результатов проводимых научных и экспериментальных исследований. В этой связи представляется интересным отметить работы ИПИ РАН по разработке **региональных сетей передачи данных** [28].

Так, в конце 1991 г. И. А. Мизин получил предложение от Государственного комитета по науке и технике построить сеть передачи данных на территории Псковской области. Заказчиком сети была определена администрация области. Руководителем этой работы он назначил одного из своих заместителей доктора технических наук Ю. В. Гайковича. Примерно год ушел на решение организационных вопросов по постановке научно-исследовательских и опытно-конструкторских работ (НИОКР), а также разработку концепции Псковской региональной информационно-телекоммуникационной сети (РИТС).

Принципиальным положением концепции стала ориентация на использование отечественных аппаратно-программных средств. При этом ИПИ РАН уже имел некоторые заделы. Так, И. А. Мизин сформулировал общий замысел отечественного коммутатора на основе протокола X.121 (X.25, X.28). Решение технических вопросов было поручено отделу А. В. Комарова (В. И. Гребенщиков, А. Б. Кулаков, А. М. Зараев и др.), разработка ПО — отделу В. А. Козмидиади (А. Н. Макаров, Л. М. Козлова, М. Я. Вайнштейн) и лаборатории С. М. Рождественского (В. А. Кондрашев, К. И. Волович, Т. В. Бондаренко). При этом за основу был принят периферийный восьмипортовый процессор графической станции «Беста», серийно выпускаемой ООО при заводе ЗИЛ [29].

Небольшой, но важный вклад в разработку коммутатора сделал Орловский филиал ИПИ РАН (О. П. Архипов), который по поручению И. А. Мизина в кратчайшие сроки изготовил конструктив для коммутатора.

С января по май 1995 г. под руководством Ю. В. Гайковича (А. Н. Макаров, В. А. Козмидиади, В. И. Гребенщиков, А. Б. Кулаков и др.) были проведены работы по оснащению 24 районных центров Псковской области ЦКП, пакетными адаптерами данных и модемами. Для управления процессом настройки сети на стенде ИПИ РАН был создан дублирующий пункт управления, соединенный каналом связи с опорной сетью в Пскове. Такое системное решение позволило оперативно устранить несколько программных ошибок. Тридцатого июня 1995 г. был подписан акт о введении Псковской РИТС в эксплуатацию [29]. Сеть предоставляла такие услуги, как электронная почта, передача файлов, телекс, телефакс.

Фактически в России в период лавинообразного внедрения импортных средств впервые была создана отечественная ИТКС. К сожалению, эти работы не получили дальнейшего продолжения: предложения по созданию аналогичных региональных СПД в Воронежской, Новгородской, Курской областях, а также в Республике Беларусь не были приняты по разным причинам, включая отсутствие финансирования.

Следует кратко упомянуть еще ряд **прикладных информационных технологий**, разработанных в ИПИ РАН под руководством И. А. Мизина и реализованных в различных сферах российского общества.

Технологии информатизации районных отделений милиции. В ИПИ РАН под руководством И. П. Кузнецова (С. О. Новиков, А. В. Губин, М. М. Шарнин, В. П. Кузнецов) был создан и апробирован в Октябрьском РОВД г. Москвы комплекс новых информационных технологий, который обеспечивал информационную поддержку решения некоторых задач оперативно-розыскной деятельности. Комплекс был реализован на персональных компьютерах типа IBM PC AT и включал в себя подсистемы локальных БД дактилоскопической информации для уровня РОВД, поиска информации в этих базах для идентификации личности правонарушителей (АДИС «УЗОР»), а также формирования БД об оперативных сводках о правонарушениях в данном районе города (система «Криминал»). Реализация проекта осуществлялась в тесном контакте с сотрудниками Октябрьского РОВД [30].

Многотерминальный комплекс информатизации объектов социальной сферы. Такой комплекс был создан в 1993 г. специалистами Казанского филиала ИПИ РАН (И. З. Насыров, Р. И. Хасанов) на базе персональной ЭВМ IBM PC AT [31]. В компьютер были встроены мультиплексоры, которые позволяли подключать к нему до 16 терминалов различного типа от малых и больших ЭВМ. Было разработано ПО для работы в мультипрограммном режиме, что позволяло использовать этот комплекс на различных объектах социальной сферы: в узлах связи, службах занятости, банках, медицинских учреждениях. На основе этой разработки был создан КСИ для узлов связи, а также для районных отделов Министерства социальной защиты населения России.

Технологии создания цифровых архивов текстовой информации. Разработанные в ИПИ РАН технологии обработки изображений оказались полезными при создании электронных архивов текстовых документов (В. С. Кот, С. О. Новиков, А. В. Губин, В. И. Дымков). Они позволяли существенно повысить качество хранимых изображений этих документов, которые могли быть как печатными, так и рукописными. На основе этих технологий был создан КСА для формирования электронного архива рукописных текстов русского языка в Российской государственной библиотеке (К. К. Колин, И. М. Зацман, Г. А. Левина). Этот проект был выполнен в рамках программы ЮНЕСКО по обеспечению сохранности культурных ценностей мирового уровня. Его реализация позволила не только обеспечить сохранность этих рукописей, изображения которых были записаны на оптических дисках CD-ROM, но также создать возможности для удаленного доступа к ним заинтересованных российских и зарубежных специалистов.

Технологии для пользователей сети Интернет. Глубоко понимая стратегическую значимость сети Интернет для развития общества, И. А. Мизин создал в структуре ИПИ РАН специальное подразделение для проведения исследований в этой области. Его руководителем стал И. М. Адамович. В результате исследо-

ваний (И. М. Адамович, О. И. Волков) были разработаны российские средства поддержки работы пользователей сети Интернет, которые учитывали передовые мировые тенденции развития сетевых технологий, а также создан российский интернет-браузер «Ариадна», который позволял пользователям работать в англоязычном варианте (под управлением Windows NT) и в русскоязычном варианте (под управлением Windows-95) [32, 33]. Расширенная версия этого браузера обеспечивала возможность передачи по сети Интернет речевой информации. Для этого было разработано программное приложение AMSD EASY TALK, которое позволяло осуществлять качественную голосовую связь между удаленными абонентами на больших расстояниях. Для России с ее огромной территорией эта технология позволяла существенно снизить стоимость телефонных переговоров между абонентами сети Интернет.

Технологии информатизации деятельности работников офисов. В условиях комплексной информатизации работы офисов различных учреждений центральное место занимает электронное хранилище большого числа текстовых документов. Разработку ПО системы под названием «Электронный архив» в 1990 г. осуществил отдел специалистов ИПИ РАН под руководством В. Д. Праченко совместно с сотрудниками Бердянского филиала ИПИ АН (К. К. Колин, И. М. Зацман, Л. Г. Гордон, О. А. Яворский и др.).

Результат этой разработки стал известен специалистам крупной американской компании Wise Technology, которая предложила И. А. Мизину создать по этой проблематике совместное предприятие. Деловые контакты с представителями этой фирмы показали, что уровень квалификации специалистов ИПИ РАН был достаточно высоким.

Российско-американский проект по модернизации терминалов серверов в системах управления и обработки данных. В рамках этого проекта (руководитель — заведующий сектором кандидат технических наук И. М. Зацман) было разработано ПО для терминалов фирмы Wise Technology, обеспечившее их применение в российских условиях [34]. Внедрить результаты этого проекта планировалось сначала в аэропортах г. Москвы, а затем и в других городах России.

Учитывая масштабы проекта, руководство фирмы Wise Technology предложило И. А. Мизину рассмотреть вопрос о строительстве в России завода этой фирмы для промышленного производства модернизированных терминалов в необходимых объемах. К сожалению, известные события, связанные с распадом Советского Союза, помешали реализации этого важного проекта.

Информационно-аналитическая система для токсикологических центров России. Эта система была создана совместными усилиями специалистов ИПИ РАН, ВЦ РАН и Информационно-консультационного токсикологического центра Московского института скорой помощи им. Склифосовского. Инициатором проекта стал директор этого Центра профессор В. Н. Дагаев, который в течение 20 лет делал выписки из историй болезни пациентов своего Института и, по сути, создал уникальную базу знаний в области диагностики и лечения острых химических отравлений.

Специалисты ВЦ РАН под руководством доктора физико-математических наук К. В. Рудакова разработали цифровое представление этой базы в памяти одной из моделей ЕС ЭВМ. Оно стало основой для создания информационно-аналитической экспертной системы для токсикологических центров России, которую разработали специалисты ИПИ РАН (В. В. Пузанов). Система обеспечивала оперативный поиск в базе знаний текстовой информации историй болезни тех пациентов, у которых ранее наблюдались аналогичные симптомы острого химического отравления. Экспертная система была успешно апробирована в деятельности Токсикологического центра Московского института скорой помощи и планировалась для распространения в других аналогичных центрах крупных городов России.

Рассмотренные выше проекты показывают тот широкий диапазон тематики исследований и разработок, которые были развернуты в ИПИ РАН в интересах развития информационного общества в России.

Как видим, И. А. Мизиным вместе с научными коллективами ИПИ РАН были сформулированы основные принципы и системотехнические подходы к решению широчайшего класса автоматизированных систем органов государственной власти РФ.

За 10 лет руководства И. А. Мизиным научные коллективы ИПИ РАН получили бесценный опыт и накопили огромный научно-практический задел в области создания информационных систем. В 1999 г. Институт возглавил И. А. Соколов (в 2003 г. избран членом-корреспондентом, в 2008 г. — академиком РАН), который несколько лет руководил системным отделом ИПИ РАН. Игорь Анатольевич стал последовательным преемником системных и методических подходов, сформированных в ИПИ РАН при И. А. Мизине.

В период 1999–2024 гг. в рамках направления «Информационные, управляющие и телекоммуникационные системы» под общим руководством академика И. А. Соколова коллективами ИПИ РАН (с 2015 г. — ФИЦ ИУ РАН) был выполнен целый ряд важнейших проектов по созданию крупных территориально распределенных ИТКС различного назначения в интересах органов государственной власти, в том числе в защищенном исполнении.

Отметим кратко два комплекса работ, отмеченных премией Правительства РФ.

В период 2001–2014 гг. научные коллективы ИПИ РАН выполнили комплекс НИОКР, направленных на развитие **ИТКС Банка России** (главный конструктор — доктор технических наук В. И. Будзко). В этой работе активное участие приняли сотрудники ИПИ РАН В. Г. Беленков, И. И. Быстров, П. А. Кейер, С. В. Борохов, Э. В. Киселев, Ю. Н. Мирошников, В. П. Бабурина, А. И. Надеждина и др. Разработаны концепции развития ИТКС Банка России на период до 2008, затем до 2013 и 2018 гг., которые были приняты в качестве основы технической политики. Обоснована технология перехода от существовавших 78 региональных центров обработки информации (ЦОИ) к трем высокопроизводительным территориально распределенным коллективным ката-

строфоустойчивым ЦОИ. Подготовлены технико-экономические обоснования их структуры и конфигурации (на начальной стадии и при развитии). Создана территориально распределенная информационная инфраструктура, обеспечивающая бесперебойную работу автоматизированных систем банковских расчетов в условиях катастроф и других непредвиденных обстоятельств при сохранении необходимого уровня информационной безопасности в случае перевода обработки с одной площадки на другую. Одновременно проводился комплекс работ по развитию телекоммуникационной составляющей, средств обеспечения информационной безопасности, управления функционированием и эксплуатацией системы коллективной обработки информации. В 2011 г. сотрудники ИПИ РАН (В. И. Будзко, И. И. Быстров, В. Г. Беленков) в составе коллектива разработчиков ИТКС Банка России были удостоены премии Правительства РФ в области науки и техники.

В период 2006–2021 гг. ИПИ РАН выполнил комплекс НИОКР по созданию, развертыванию и развитию **информационно-аналитической системы сбора информации от большого числа территориально распределенных источников** (главный конструктор — А. А. Зацаринный, ответственный исполнитель — В. И. Сеницын, исполнители — И. М. Адамович, Т. К. Бирюкова, М. М. Гершкович, М. Ю. Заикин, В. С. Долгополов, Л. М. Козлова). Система на протяжении многих лет успешно выполняет важнейшие задачи по обеспечению безопасности страны, в том числе в период ряда социально значимых международных мероприятий. В 2021 г. коллектив исполнителей, включая сотрудников ФИЦ ИУ РАН (А. А. Зацаринный, В. И. Сеницын, Т. К. Бирюкова, М. М. Гершкович), удостоен премии Правительства РФ в области науки и техники.

Таким образом, научная школа академика И. А. Мизина в части методологических подходов и базовых системотехнических решений по созданию территориально распределенных автоматизированных информационных систем на протяжении многих лет активно поддерживается коллективом ИПИ РАН, а с 2015 г. — ФИЦ ИУ РАН.

4 Вклад И. А. Мизина в развитие информационного общества

Для решения стратегически важных проблем развития информационного общества в России особую значимость имеют разработанные И. А. Мизиным **концептуальные основы создания единого информационного пространства страны**. Представление о значимости этого результата дают основные положения его статьи [22], в которой показана ключевая роль в формировании информационного общества телекоммуникационных технологий, а также отмечена необходимость международного сотрудничества для построения информационного общества. Сформулирован концептуальный тезис: «Приближающийся XXI век можно смело назвать веком “информационного сообщества”». Эти прогнозы оказались справедливыми и в настоящее время уже осуществляются на практике.

Исследования проблем развития информационного общества И. А. Мизин осуществлял совместно с сотрудниками ИПИ РАН: В. Н. Захаровым, И. М. Зацманом, К. К. Колиным, И. П. Кузнецовым, Н. А. Левиным, В. Д. Праченко, Е. Н. Филиновым и другими специалистами. При этом он активно взаимодействовал с ведущими учеными РАН: Е. П. Велиховым (академиком-секретарем ОИВТА РАН), В. С. Бурцевым, Ю. В. Гуляевым, А. А. Дородницыным, С. В. Емельяновым, а также с профильными зарубежными специалистами из Германии, Испании, Франции и США.

Рассмотрим несколько направлений деятельности в этой области.

В 1996 г. в пленарном докладе на встрече нобелевских лауреатов в Санкт-Петербурге И. А. Мизин представил результаты исследований в области **научно-технологических основ становления глобального информационного общества**. Доклад был подготовлен по просьбе академика Ж. И. Алферова и вызвал большой интерес участников этой встречи. Его детальное обсуждение прошло на заседании одной из тематических секций этого научного мероприятия.

Аналогичный доклад был сделан И. А. Мизиным и на расширенном бюро ОИВТА РАН, посвященном проблематике развития информационного общества в марте 1995 г. Было рекомендовано провести по этой проблематике комплекс исследований и разработать для этого целевую программу Президиума РАН. Однако недостаток финансирования в тот период времени не позволил осуществить эту рекомендацию.

В подготовке этих докладов приняли участие сотрудники ИПИ РАН И. М. Адамович, С. Н. Гринченко, В. Н. Захаров, К. К. Колин, И. П. Кузнецов, И. Н. Синицын и Е. Н. Филинов.

Российская концепция информатизации образования. Развитие информационного общества требует нового уровня информационной культуры населения страны, который необходим для эффективного использования национальных и мировых информационных ресурсов, а также тех новых возможностей, которые создают современные информационные системы и технологии. А это, в свою очередь, вызывает необходимость адекватной коррекции содержания и методологии системы образования на всех ее уровнях, начиная с дошкольной подготовки.

Игорь Александрович Мизин глубоко понимал сущность этой стратегически важной проблемы и уделял ей значительное внимание. В ряде работ [35, 36], прежде всего в докладе на пленарном заседании II Международного конгресса ЮНЕСКО «Образование и информатика» (июнь 1996 г.), он представил основные положения стратегии России в области создания и развития национальной ИТКС для науки и образования с одновременным развитием таких систем в регионах нашей страны и созданием в них банков данных компьютерной информации [37].

Нужно отметить, что упомянутый выше сборник трудов ИПИ РАН («Системы и средства информатики». Вып. 8. — М.: Наука, 1996) [35] стал вторым тематическим выпуском, который был полностью посвящен проблемам

информатизации образования. Первый такой выпуск был издан в 1988 г. и посвящен в основном проблемам формирования *компьютерной грамотности общества*. А главной темой второго выпуска стала проблема формирования *новой информационной культуры общества*, которая рассматривалась как одна из важнейших целей системы образования в условиях становления глобального информационного общества.

Пленарный доклад И. А. Мизина на Конгрессе ЮНЕСКО под названием «Состояние и перспективы развития информационных телекоммуникационных технологий для сферы образования и науки» имел концептуальный характер [38]. Он был посвящен анализу состояния, проблем и перспектив развития этих технологий как основы для создания в XXI в. глобального информационного пространства нашей планеты. При этом И. А. Мизин изложил свое понимание термина «информационная инфраструктура» как такой совокупности информационных ресурсов и программно-аппаратных средств вычислительной и телекоммуникационной техники, информационных технологий и телекоммуникационных сетей, которая будет использоваться во всех сферах информационного общества.

Этот доклад стал важным событием в работе Конгресса ЮНЕСКО, а его текст включен в итоговые материалы Конгресса. В его подготовке приняли участие К. К. Колин и В. Н. Захаров.

Информатика в системе опережающего образования. Второй Международный конгресс ЮНЕСКО «Образование и информатика» состоялся в России в 1996 г. на базе Московского государственного университета им. М. В. Ломоносова. Его участниками стали делегации из более 100 стран мира, многие из которых, включая Россию, представили национальные доклады о политике в области образования на основе использования новых знаний и перспективных информационных технологий.

Конгресс стал крупным международным событием для мирового научно-образовательного сообщества. Президентом Конгресса был избран академик Российской академии образования (РАО) В. Г. Кинелев, который в тот период времени был заместителем Председателя правительства РФ и Председателем Государственного комитета по высшему образованию. В своем пленарном докладе на этом Конгрессе «*Цивилизация и образование*» он задал высокую научную планку обсуждения его основной проблематики, указав на необходимость концентрации внимания не на инструментально-технологических, а главным образом на содержательных ее аспектах. Именно такая направленность и отличала доклады российских участников этого Конгресса: К. К. Колина, С. А. Христочевского, И. В. Роберт, И. И. Халеевой, А. Д. Урсула. Так, в докладе «*Информатика в системе опережающего образования*» была представлена российская концепция использования достижений информатики для опережающего развития системы образования на фоне других факторов социально-экономического и культурного развития общества [38].

Эта концепция была разработана в ИПИ РАН под руководством И. А. Мизина (К. К. Колин, С. А. Христочевский), а ее основные положения, по предложе-

нию И. А. Мизина, были включены в Национальный доклад России Конгрессу ЮНЕСКО [39].

Для практической реализации концепции предлагалось выполнить под эгидой ЮНЕСКО в период 1997–2000 гг. международный проект «Информатика-2000». Цель проекта состояла в том, чтобы радикальным образом изменить содержание образования на всех его уровнях с учетом перспектив развития информационного общества [40]. Концепция была одобрена участниками Конгресса и включена в качестве рекомендации в его итоговые документы.

Основная идея этого проекта состояла в том, чтобы разработать и внедрить на всех уровнях системы образования новый общеобразовательный курс «Фундаментальные основы информатики». Он включал следующие разделы:

- теоретические основы информатики;
- технические и программные средства информатизации;
- информационные технологии;
- социальная информатика.

Напомним, что в тот период времени информатика рассматривалась как техническая дисциплина, в которой изучались способы хранения, обработки и передачи данных при помощи компьютерных систем. А новый курс был ориентирован на существенно более широкую проблематику становления информационного общества.

В 1997 г. в Россию приехала группа экспертов ЮНЕСКО, которая рекомендовала начать реализацию этого проекта сначала в России, а затем, после практической апробации его результатов, распространить его на другие страны. С этой целью И. А. Мизиным и ректором Московского экономико-статистического института (МЭСИ) В. П. Тихомировым было принято решение о создании совместного Научно-образовательного центра «*Фундаментальные основы информатики*». Для его развития был получен грант Российского фонда фундаментальных исследований, который позволил разработать необходимые научно-методические материалы по этому проекту.

Серьезную поддержку этому проекту оказал академик РАО В. Г. Кинелев, который стал директором нового Института ЮНЕСКО по информационным технологиям в образовании. Этот Институт был создан в Москве сразу же по окончании Конгресса ЮНЕСКО. Решение об этом принял директор ЮНЕСКО Фредерико Майор, несмотря на противодействие со стороны некоторых представителей руководящих органов этой организации из стран Запада.

Апробация результатов проекта в системе образования была осуществлена в МЭСИ, в Московском социальном университете, а также в Академии государственной службы при Президенте РФ. Она оказалась успешной и в последующие годы получила свое развитие в ряде других учреждений высшего и среднего образования нашей страны [41].

5 О книге «И. А. Мизин — ученый, конструктор, человек»

В апреле 2009 г. на заседании Ученого совета ИПИ РАН было принято решение о подготовке монографии к 75-летию Мизина. Сформирована рабочая группа под руководством академика И. А. Соколова в составе: А. А. Зацаринный, В. Н. Захаров, И. Н. Сеницын и А. И. Темнов.

В авторский коллектив монографии рабочая группа постаралась привлечь соратников Игоря Александровича по разработке СОД, представителей заказчика, научных и промышленных организаций. Так появились статьи сотрудников НИИАА Г. К. Храмешина, Г. А. Оганяна, А. В. Тамошинского, А. И. Воронина, Л. Б. Щукина, Г. Г. Карпова, В. И. Петрова. Заметки-воспоминания написал академик А. П. Кулешов, в 1980-е гг. — главный инженер НТЦ СПД, один из ближайших соратников И. А. Мизина

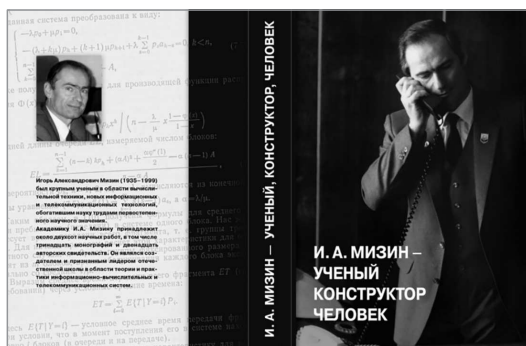
Валентин Иванович Мирошников, Главный конструктор СПД ВМФ, в своей статье рассказал не только о личном общении с Мизиным, но также о роли маршала С. Ф. Ахромеева в создании КСБУ и СОД.

Интересные заметки написал директор ЦИТИС генерал армии А. В. Старовойтов (директор ФАПСи в 1990-е гг.), с которым И. А. Мизин много взаимодействовал по созданию и модернизации информационных систем в интересах органов государственной власти.

Активное участие в подготовке книги к изданию приняли дочери И. А. Мизина. Семья Игоря Александровича — вдова Нина Федоровна, дочери и сын — написали о нем теплые заметки, а также предоставили ряд интересных материалов из семейного архива (фото, рисунки, документы).

Восьмого апреля 2010 г. ИПИ РАН организовал и провел презентацию книги в Доме ученых на Пречистенке.

На это мероприятие собралось много людей, близких И. А. Мизину: сотрудники НИИАА (бывшие и настоящие) Л. Б. Щукин, А. П. Кулешов, А. В. Ермо-



Книга «И. А. Мизин — ученый, конструктор, человек», изданная в 2010 г.



Дом ученых на Пречистенке, 8 апреля 2010 г., презентация книги о И. А. Мизине, изданной к его 75-летию. Слева направо: А. А. Водолаженко, А. А. Зацаринный, А. М. Растрелин, В. М. Якунин, А. В. Тамошинский, Ю. Б. Зубарев, С. М. Чудинов, А. П. Реутов

ленко, Л. Б. Третьяков, В. Б. Коротаев, Г. А. Оганян, Г. Г. Карпов, А. В. Бернштейн, Г. А. Кабатянский, В. И. Мирошников (из Санкт-Петербурга), Р. В. Атоян (из Еревана), В. Б. Бетелин (НИИСИ РАН), Ю. Б. Зубарев (НИИ Радио), А. П. Реутов (замминистра МРП СССР), Ю. М. Залогин (последний начальник связи, с которым работал Мизин), Н. И. Ильин (УИС ФСО России), А. А. Водолаженко (ЦОИ «Энергия» ФСО России), В. М. Якунин (начальник 16 ЦНИИИ МО).

В презентации книги приняли участие ведущие сотрудники ИПИ РАН: И. Н. Сеницын, В. И. Сеницын, В. А. Кондрашев, Э. Р. Корепанов. Из Орла приехали директор филиала ИПИ РАН О. П. Архипов и представитель Академии связи ФСО Ю. П. Остриков.

На презентации с докладами выступили И. А. Соколов, А. А. Зацаринный, А. П. Кулешов, В. И. Мирошников, Р. В. Атоян.

Главное, что книга всем понравилась.

6 Заключение

Кроме научной и научно-практической деятельности по тематике создания и внедрения информационных и телекоммуникационных систем, непосредственного руководства большими коллективами И. А. Мизин всегда вел большую научно-организационную работу. В советские годы он был председателем секции «Теория передачи и обработки информации» Научного совета по комплексной проблеме «Кибернетика» АН СССР, членом Комиссии по вычислительным центрам коллективного пользования и сетям ЭВМ при Президиуме АН СССР, членом секции Межправительственной комиссии по вычислительной технике,

членом Совета руководителей Академсети, членом советской рабочей группы Международной организации по стандартизации. В годы работы в ИПИ РАН И. А. Мизин был председателем Научного совета РАН по информационно-вычислительным сетям, членом Научного совета по выставкам РАН, членом Координационного совета по информатизации Администрации Президента Российской Федерации, членом Президиума научно-технического совета по программе «Информатизация России», председателем Головного совета главных конструкторов АСУ ВС РФ, членом научных советов при Совете безопасности РФ, Совете обороны РФ, Минэкономки России, членом Президиума НТС г. Москвы.

Игорь Александрович Мизин активно участвовал в организации многих отечественных и международных семинаров, совещаний и конференций. Он был главным редактором журнала «Микропроцессорные средства и системы» и ежегодника ИПИ РАН «Системы и средства информатики», членом редколлегий еще ряда журналов. С 1996 г. и до последних дней И. А. Мизин активно участвовал в работе ОИВТА РАН в качестве члена бюро Отделения.

Игорь Александрович Мизин много внимания уделял подготовке научных и инженерных кадров. Он был заведующим базовыми кафедрами в Московском государственном университете радиотехники, электроники и автоматики и в Московском техническом университете связи и информатики, председателем двух специализированных советов по защите диссертаций. Подготовил 9 докторов и 15 кандидатов наук.

Большая научная эрудиция, инженерные знания и высокие человеческие качества позволяли И. А. Мизину создавать в науке и промышленности успешно работающие большие коллективы специалистов, объединять вокруг себя талантливую молодежь.

Научная, научно-организационная и педагогическая деятельность И. А. Мизина была по достоинству оценена нашей Родиной. Он был лауреатом Ленинской и Государственной премий, награжден орденом Трудового Красного Знамени, орденом «За заслуги перед Отечеством» 4-й степени, многими медалями.

Игорь Александрович Мизин умер в расцвете творческих сил 8 сентября 1999 г. после тяжелой болезни. Похоронен в Москве на Востряковском кладбище.

Память об И. А. Мизине поддерживается коллективом ИПИ РАН. К 75-летию со дня рождения была выпущена монография [3], в которой всесторонне представлены различные виды деятельности И. А. Мизина, но также много внимания уделено его человеческим качествам. О его вкладе в создание СОД подробно рассказано в монографии [6].

Авторы статьи выражают благодарность за предоставленные материалы, а также замечания и рекомендации В. И. Будзко, В. И. Гребенчикову, А. М. Зараеву, С. В. Козлову, В. А. Кондрашеву, Э. Р. Корепанову, И. Н. Синицыну, В. И. Синицыну, А. Н. Порошаю.

Литература

1. Соколов И. А., Зацаринный А. А. Игорь Александрович Мизин // Военно-промышленная комиссия. 60 лет на страже Родины. — М.: Оружие и технологии, 2017. С. 921–926. EDN: YPDHIV.
2. Автоматизация управления. Наш путь: к 50-летию НИИ автоматической аппаратуры им. акад. В. С. Семенихина / Под ред. А. А. Антонова. — М.: НИИАА, 2006. 210 с.
3. И. А. Мизин — ученый, конструктор, человек / Под ред. акад. И. А. Соколова. — М.: ИПИ РАН, 2010. 320 с. doi: 10.14357/08696527150114. EDN: UCMGXB.
4. Соколов И. А., Зацаринный А. А., Захаров В. Н. О вкладе академика И. А. Мизина в теорию и практику создания отечественных информационно-телекоммуникационных систем: к 80-летию со дня рождения // Системы и средства информатики, 2015. Т. 25. № 1. С. 213–229. doi: 10.14357/08696527150114. EDN: UCMGXB.
5. Соколов И. А., Зацаринный А. А., Захаров В. Н. У истоков Российской информатики // Системы и средства информатики, 2020. Т. 30. № 2. С. 195–202. EDN: TZKECW.
6. Зацаринный А. А. Система обмена данными: люди, события, факты. — М.: ТОРУС ПРЕСС, 2021. 516 с. doi: 10.30826/94588-289-8. EDN: JWYJOJ.
7. Петухов В. Е., Жуков В. А., Колодцев Ю., Москевич В. В. К истории становления «ядерной кнопки» России. — СПб.: СПбГПУ, 2003. 488 с.
8. Щукин Л. Б. Игорь Мизин: системное мышление, новаторство, целеустремленность // И. А. Мизин — ученый, конструктор, человек / Под ред. И. А. Соколова. — М.: ИПИ РАН, 2010. С. 136–141.
9. Петров В. И. Как хороши, прекрасны были годы! // И. А. Мизин — ученый, конструктор, человек / Под ред. И. А. Соколова. — М.: ИПИ РАН, 2010. С. 238–239.
10. Мизин И. А., Уринсон Л. С., Храмешин Г. К. Передача информации в сетях с коммутацией сообщений. — М.: Связь, 1972. 319 с.; 2-е изд., 1977. 328 с.
11. Cerf V., Kahn R. A protocol for packet network intercommunication // IEEE T. Commun., 1974. Vol. 22. Iss. 5. P. 637–648. doi: 10.1109/tcom.1974.1092259.
12. Kleinrock L. Queueing systems: Volume I — theory. — New York, NY, USA: Wiley Interscience, 1975. 417 p.
13. Сети ЭВМ / Под ред. В. М. Глушкова. — М.: Связь, 1977. 280 с.
14. Отечественная электронная вычислительная техника: биографическая энциклопедия / Сост. С. А. Муравьев. — М.: Столичная энциклопедия, 2014. 400 с. [Матюхин — с. 141; Брук — с. 29.]
15. Зацаринный А. А. Теория и практика системного подхода: к 70-летию профессора В. А. Богатырева // ВКС. CONNECT, 2007. № 4. С. 166–169.
16. Халин В. Г., Зацаринный А. А., Геков В. В. и др. Об опыте создания и развития АСУ ВС СССР и РФ / Под ред. В. Г. Халина. — М.: АдамантЪ, 2019. Т. 1. 647 с.; Т. 2. 264 с.; Т. 3. 231 с.
17. Зацаринный А. А. Академик Мизин И. А. — военная наука и практика // И. А. Мизин — ученый, конструктор, человек / Под ред. И. А. Соколова. — М.: ИПИ РАН, 2010. С. 96–128.

18. Зацаринный А. А. Военный институт связи на службе Родине: к 100-летию создания. — М.: ТОРУС ПРЕСС, 2023. 530 с. doi: 10.30826/94588-315-4. EDN: SRYTQM.
19. Ушаков И. А. Воспоминания // И. А. Мизин — ученый, конструктор, человек / Под ред. И. А. Соколова. — М.: ИПИ РАН, 2010. С. 213–215.
20. Мизин И. А., Богатырев В. А., Кулешов А. П. Сети коммутации пакетов. — М.: Радио и связь, 1986. 408 с.
21. Протоколы информационно-вычислительных сетей: Справочник / Под ред. И. А. Мизина, А. П. Кулешова. — М.: Радио и связь, 1990. 504 с.
22. Мизин И. А. Телекоммуникационные технологии: состояние и перспективы развития // Электроника: наука, технология, бизнес, 1998. № 1. С. 13–18. EDN: WTJZZT.
23. Мизин И. А. Современное состояние проблематики интегрированных информационно-телекоммуникационных систем и сетей // Системы и средства информатики, 1999. Вып. 9. С. 11–23.
24. Антонов С. В., Соколов И. А., Шоргин С. Я. Математическое моделирование как инструмент исследования и проектирования телекоммуникационных систем // Системы и средства информатики, 1999. Вып. 9. С. 34–44.
25. Соколов И. А., Зацаринный А. А., Антонов С. В., Печинкин А. П., Шоргин С. Я. Модели функционирования информационно-телекоммуникационных систем (подраздел 3.3) // Методическое руководство по оценке качества функционирования информационных систем (в контексте стандарта ГОСТ РВ 51987-2002 Информационная технология. Комплекс стандартов на автоматизированные системы. Требования и показатели качества функционирования информационных систем. Общие положения). — М.: 3 ЦНИИ МО РФ, 2003. С. 157–215.
26. Соколов И. А., Зацаринный А. А., Печинкин А. П., Антонов С. В., Шоргин С. Я., Душин Ю. А., Лызлова И. В. Комплекс программно-математических средств моделирования информационно-телекоммуникационных систем // Системы и средства информатики, 2006. Вып. 16. С. 4–31. EDN: KZUWED.
27. Рейтинг популярности СУБД DB-Engines. 2024. <https://db-engines.com/en/ranking>.
28. Гайкович Ю. В. Опыт создания и концепция построения региональных информационно-телекоммуникационных систем в регионах России // Системы и средства информатики, 1999. Вып. 9. С. 23–33.
29. Гайкович Ю. В. Региональные телекоммуникации от Мизина // И. А. Мизин — ученый, конструктор, человек / Под ред. И. А. Соколова. — М.: ИПИ РАН, 2010. С. 193–196.
30. Кузнецов И. П., Дьяченко И. М., Шарнин М. М. Система по раскрытию организованных преступных формирований СПРУТ-93 // Социальная информатика-93. — М.: АЕН РФ, ИПИ РАН, Институт молодежи, 1993. С. 76–81.
31. Колин К. К., Кокотов В. М., Насыров И. Э., Тимофеев В. И., Хасонов Р. И., Заритов Э. Ф., Гилязетдинов И. М., Крылов Е. Е., Стариков А. Л. Применение многотерминальных комплексов на базе ПЭВМ для автоматизации обработки данных в предприятиях социальной сферы // Социальная информатика-93. — М.: АЕН РФ, ИПИ РАН, Институт молодежи, 1993. С. 152–155.
32. Аллатов В. Нить Ариадны // Компьютерра, 11 ноября 1996. № 42. С. 51.

33. *Адамович И. А.* Российские средства поддержки технологий Интернет // Социальная информатика-98. — М.: ИПИ РАН, Институт молодежи, 1998. С. 116–121.
34. *Колин К. К., Зацман И. М., Левина Г. А.* Новые информационные технологии для сохранения и эффективного использования рукописной картотеки русского языка XI–XVII вв. на CD-ROM // Социальная информатика-95. — М.: Институт молодежи, 1995. С. 80–82. EDN: UGKFRB.
35. *Мизин И. А., Колин К. К.* Информационные и телекоммуникационные технологии в системе образования России // Системы и средства информатики, 1996. Вып. 8. С. 1–13. EDN: VONAQZ.
36. *Мизин И. А.* Состояние и перспективы развития информационных и телекоммуникационных технологий: Доклад на II Международном конгрессе ЮНЕСКО «Образование и информатика», 1996 // Информационное общество, 1996. Вып. 5. С. 3–22.
37. *Мизин И. А., Киселев Э. В., Соколов И. А., Шоргин С. Я.* Некоторые проблемы создания единой информационно-телекоммуникационной системы общенационального масштаба как основы информатизации сферы образования в России // Системы и средства информатики, 1996. Вып. 8. С. 114–124.
38. *Колин К. К.* Информатика в системе опережающего образования // Вестник Российского общества информатики и вычислительной техники, 1996. № 3. С. 19–39. EDN: WLNXTV.
39. Политика в области образования и новые информационные технологии: Национальный доклад Российской Федерации // Труды II Международного конгресса ЮНЕСКО «Образование и информатика». — М., 1996. 34 с.
40. *Колин К. К.* «Информатика-2000» — международный проект для системы образования // Международное сотрудничество, 1997. № 1. С. 17–18. EDN: WIXSSZ.
41. *Колин К. К.* Образование для информационного общества: проблемы и приоритеты // Информационное общество, 2022. № 5. С. 16–34. doi: 10.52605/16059921.2022.05.16. EDN: CIQJJS.

Поступила в редакцию 09.01.2025

Принята к публикации 15.02.2025

THE ROLE OF THE ACADEMICIAN I. A. MIZIN IN INFORMATION TECHNOLOGY DEVELOPMENT AND APPLICATION FOR GOVERNMENT NEEDS: ON THE 90TH ANNIVERSARY OF HIS BIRTH

I. A. Sokolov, A. A. Zatsarinny, V. N. Zakharov, and K. K. Kolin

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

Abstract: The paper is dedicated to the 90th anniversary of the academician I. A. Mizin — the esteemed scientist in fields of telecommunications, computer

science, and information technologies. Here, one focuses on the most significant results he obtained during the 40 years of his scientific work. While working as the chief designer at the Research Institute of Automatic Equipment, I. A. Mizin ensured successful completion of a set of works on the development and implementation of the first domestic geographically distributed data exchange system with packet switching by his team for the purposes of automated control systems of strategic importance. Under his leadership, the research teams of the Institute of Informatics Problems of the Russian Academy of Sciences (RAS) substantiated the choice of information technologies for the creation and development of information networks that comply with the requirements for information security in the interests of the governmental authorities. Igor A. Mizin was a member of the Bureau of the Department of Informatics, Computer Engineering, and Automation of the RAS, was the founder and recognized leader of the domestic scientific school in the theory and practice of information computing and telecommunications systems, and was actively engaged in research and organizational work. His contribution to the formation and development of the Russian information society is being revealed.

Keywords: chief designer; academician; information and telecommunications networks; information technologies; data exchange system; data switching center; code message switching; data transfer protocols; information security; communication channel; governmental authorities

DOI: 10.14357/08696527250101

EDN: FWHKEO

References

1. Sokolov, I. A., and A. A. Zatsarinny. 2017. Igor' Aleksandrovich Mizin [Igor' Aleksandrovich Mizin]. *Voyenno-promyshlennaya komissiya. 60 let na strazhe Rodiny* [Military-industrial commission. 60 years guarding the Homeland]. Moscow: Oruzhie i tekhnologii. 921–926. EDN: YPDHIV.
2. Antonov, A. A., ed. 2006. *Avtomatizatsiya upravleniya. Nash put': k 50-letiyu NII avtomaticheskoy apparatury im. akademika V. S. Semenikhina* [Control automation. Our path: To the 50th anniversary of academician V. S. Semenikhin Research Institute of Automated Apparature]. Moscow: NIIAA. 210 p.
3. Sokolov, I. A., ed. 2010. *Igor' Aleksandrovich Mizin — uchenyy, konstruktor, chelovek* [Igor' Aleksandrovich Mizin as a scientist, a designer, and a man]. Moscow: IPI RAS. 320 p. EDN: BZRSHV.
4. Sokolov, I. A., A. A. Zatsarinny, and V. N. Zakharov. 2015. O вкладе академика I. A. Mizina v teoriyu i praktiku sozdaniya otechestvennykh informatsionno-telekommunikatsionnykh sistem: k 80-letiyu so dnya rozhdeniya [About academician I. A. Mizin's contribution to theory and practice of domestic information-telecommunication systems creation: To the 80th anniversary]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 25(1):213–229. doi: 10.14357/08696527150114. EDN: UCMGXB.
5. Sokolov, I. A., A. A. Zatsarinny, and V. N. Zakharov. 2020. U istokov rossiyskoy informatiki [At the origins of Russian computer science]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(2):195–202. EDN: TZKECW.

6. Zatsarinny, A. A. 2021. *Sistema obmena dannymi: lyudi, sobytiya, fakty* [Data sharing system: People, events, and facts]. Moscow: TORUS PRESS. 516 p. doi: 10.30826/94588-289-8. EDN: JWYJOJ.
7. Petukhov, V., V. A. Zhukov, Yu. Kolodtsev, and V. Moskevich. 2003. *K istorii stanovleniya "yadernoy knopki" Rossii* [On the history of the formation of Russia's "nuclear button"]. Saint Petersburg: SPbGPU. 488 p.
8. Shchukin, L. B. 2010. I. A. Mizin: sistemnoe myshlenie, novatorstvo, tselestremennost' [I. A. Mizin: Systemic thinking, innovation, and sense of purpose]. *Igor' Aleksandrovich Mizin — uchenyy, konstruktor, chelovek* [Igor' Aleksandrovich Mizin as a scientist, a designer, and a man]. Ed. I. A. Sokolov. Moscow: IPI RAS. 136–141.
9. Petrov, V. I. 2010. Kak khoroshi, prekrasny byli gody [How good and wonderful the years were]. *Igor' Aleksandrovich Mizin — uchenyy, konstruktor, chelovek* [Igor' Aleksandrovich Mizin as a scientist, a designer, and a man]. Ed. I. A. Sokolov. Moscow: IPI RAS. 238–239.
10. Mizin, I. A., L. S. Urinson, and G. K. Khrameshin. 1972, 1977. *Peredacha informatsii v setyakh s kommutatsiey soobshcheniy* [Transmission of information in message switching networks]. Moscow: Svyaz. 1st ed. 319 p.; 2nd ed. 328 p.
11. Cerf, V., and R. Kahn. 1974. A protocol for packet network intercommunication. *IEEE T. Commun.* 22(5):637–648. doi: 10.1109/tcom.1974.1092259.
12. Kleinrock, L. 1975. *Queueing systems: Volume I — theory*. New York, NY: Wiley Interscience. 417 p.
13. Glushkov, V. M., ed. 1977. *Seti EVM* [Computer networks]. Moscow: Svyaz. 280 p.
14. Murav'ev, S. A., compiler. 2014. *Otechestvennaya elektronnyaya vychislitel'naya tekhnika: biograficheskaya entsiklopediya* [Domestic electronic computer technology: Biographical encyclopedia]. Ed. S. V. Khokhlov. Moscow: Capital Encyclopedia. 400 p.
15. Zatsarinny, A. A. 2007. Teoriya i praktika sistemnogo podkhoda: k 70-letiyu professora V. A. Bogatyreva [Theory and practice of a systems approach: On the 70th anniversary of professor V. A. Bogatyrev]. *VKSS. CONNECT* [VKSS. Connect] 4:166–169.
16. Khalin, V. G., A. A. Zatsarinny, V. V. Gekov, et al. 2019. *Ob opyte sozdaniya i razvitiya avtomatizirovannykh sistem upravleniya Vooruzhennykh Sil SSSR i Rossiyskoy Federatsii* [On the experience of creating and developing automated control systems of the Armed Forces of the USSR and the Russian Federation]. Ed. V. G. Khalin. Moscow: Adamant. Vol. 1. 647 p.; Vol. 2. 264 p.; Vol. 3. 231 p.
17. Zatsarinny, A. A. 2010. Akademik I. A. Mizin: Voennaya nauka i praktika [An academician I. A. Mizin: Military science and practice]. *Igor' Aleksandrovich Mizin — uchenyy, konstruktor, chelovek* [Igor' Aleksandrovich Mizin as a scientist, a designer, and a man]. Ed. I. A. Sokolov. Moscow: IPI RAS. 96–128.
18. Zatsarinny, A. A. 2023. *Voennyi institut svyazi na sluzhbe Rodine: k 100-letiyu sozdaniya* [Military institute of communications in the service of the Homeland: On the 100th anniversary of its founding]. Moscow: TORUS PRESS. 528 p. doi: 10.30826/94588-315-4. EDN: SRYTQM.
19. Ushakov, I. A. 2010. Vospominaniya [Memories]. *Igor' Aleksandrovich Mizin — uchenyy, konstruktor, chelovek* [Igor' Aleksandrovich Mizin as a scientist, a designer, and a man]. Ed. I. A. Sokolov. Moscow: IPI RAS. 213–215.
20. Mizin, I. A., V. A. Bogatyrev, and A. P. Kuleshov. 1986. *Seti kommutatsii paketov* [Packet switching networks]. Moscow: Radio i svyaz. 408 p.

21. Mizin, I. A., and A. P. Kuleshov, eds. 1990. *Protokoly informatsionno-vychislitel'nykh setey: Spravochnik* [Protocols of information and computer networks: Directory]. Moscow: Radio i svyaz. 504 p.
22. Mizin, I. A. 1998. Telekommunikatsionnye tekhnologii: sostoyanie i perspektivy razvitiya [Telecommunication technologies: State and future development]. *Elektronika: nauka, tekhnologiya, biznes* [Electronics: Science, Technology, and Business] 1:13–18. EDN: WTJZZT.
23. Mizin, I. A. 1999. Sovremennoe sostoyanie problematiki integrirrovannykh informatsionno-telekommunikatsionnykh sistem i setey [Current state of the problems of integrated information and telecommunication systems and networks]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 9:11–23.
24. Antonov, S. V., I. A. Sokolov, and S. Ya. Shorgin. 1999. Matematicheskoe modelirovaniye kak instrument issledovaniya i proektirovaniya telekommunikatsionnykh sistem [Mathematical modeling as a tool for research and design of telecommunication systems]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 9:34–44.
25. Sokolov, I. A., A. A. Zatsarinny, S. V. Antonov, A. V. Pechinkin, and S. Ya. Shorgin. 2003. Modeli funktsionirovaniya informatsionno-telekommunikatsionnykh sistem (podrazdel 3.3) [Models of functioning of information and telecommunication systems (subsection 3.3)]. *Metodicheskoe rukovodstvo po otsenke kachestva funktsionirovaniya informatsionnykh sistem (v kontekste standarta GOST RV 51987-2002 “Informatsionnaya tekhnologiya. Kompleks standartov na avtomatizirovannyye sistemy. Trebovaniya i pokazateli kachestva funktsionirovaniya informatsionnykh sistem. Obshchie polozheniya”)* [Methodological guidelines for assessing the quality of the functioning of information systems (in the context of the standard GOST RV 51987-2002 “Information technology. A set of standards for automated systems. Requirements and indicators of the quality of the functioning of information systems. General provisions”)]. Moscow: 3 TsNII MO RF. 157–215.
26. Sokolov, I. A., A. A. Zatsarinny, A. V. Pechinkin, S. V. Antonov, S. Ya. Shorgin, Yu. A. Dushin, and I. V. Lyzlova. 2006. Kompleks programmno-matematicheskikh sredstv modelirovaniya informatsionno-telekommunikatsionnykh sistem [Complex of mathematical and program tools for information and telecommunication systems modeling]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 16:4–31. EDN: KZUWED.
27. Reyting populyarnosti SUBD DB-Engines [DB-Engines DBMS popularity rating]. Available at: <https://db-engines.com/en/ranking> (accessed March 2, 2025).
28. Gaykovich, Yu. V. 1999. Opyt sozdaniya i kontseptsiya postroeniya regional'nykh informatsionno-telekommunikatsionnykh sistem v regionakh Rossii [Experience of creation and concept of construction of regional information and telecommunication systems in the regions of Russia]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 9:23–33.
29. Gaykovich, Yu. V. 2010. Regional'nye telekommunikatsii ot Mizina [Regional telecommunications by Mizin]. *Igor' Aleksandrovich Mizin — uchenyy, konstruktor, chelovek* [Igor' Aleksandrovich Mizin as a scientist, a designer, and a man]. Ed. I. A. Sokolov. Moscow: IPI RAS. 193–196.
30. Kuznetsov, I. P., I. M. D'yachenko, and M. M. Sharnin. 1993. Sistema po raskrytiyu organizovannykh prestupnykh formirovaniy SPRUT-93 [System for detecting orga-

- nized crime groups SPRUT-93]. *Sotsial'naya informatika-93* [Social Informatics-93]. Moscow: AEN RF, IPI RAS, Institute of Youth. 76–81.
31. Kolin, K. K., V. M. Kokotov, I. E. Nasyrov, V. I. Timofeev, R. I. Khasonov, E. F. Zaripov, I. M. Gilyazetdinov, E. E. Krylov, and A. L. Starikov. 1993. Prime-
nenie mnogoterminal'nykh kompleksov na baze PEVM dlya avtomatizatsii obrabotki
dannyykh v predpriyatiyakh sotsial'noy sfery [Application of PC-based multitermi-
nal complexes for automation of data processing in social enterprises]. *Sotsial'naya
informatika-93* [Social informatics-93]. Moscow: AEN RF, IPI RAS, Institute of
Youth. 152–155.
 32. Alpatov, V. 1996. Nit' Ariadny [Ariadne's thread]. *Kompyuterra* [Computerra] 42:51.
 33. Adamovich, I. A. 1998. Rossiyskie sredstva podderzhki tekhnologiy Internet [Russian
internet technology support tools]. *Sotsial'naya informatika-98* [Social informatics-98].
Moscow: IPI RAS, Institute of Youth. 116–121.
 34. Kolin, K. K., I. M. Zatsman, and G. A. Levina. 1995. Novye informatsionnye
tekhnologii dlya sokhraneniya i effektivnogo ispol'zovaniya rukopisnoy kartoteki russko-
go yazyka XI–XVII vv. na CD-ROM [New information technologies for preserving and
effectively using the handwritten card index of the Russian language of the 11th–17th
centuries on CD-ROM]. *Sotsial'naya informatika-95* [Social informatics-95]. Moscow:
Institute of Youth. 80–82. EDN: UGKFRB.
 35. Mizin, I. A., and K. K. Kolin. 1996. Informatsionnye i telekommunikatsionnye
tekhnologii v sisteme obrazovaniya Rossii [Information and telecommunication tech-
nologies in the education system of Russia]. *Sistemy i Sredstva Informatiki — Systems
and Means of Informatics* 8:1–13. EDN: VONAQZ.
 36. Mizin, I. A. 1996. Sostoyanie i perspektivy razvitiya informatsionnykh i telekom-
munikatsionnykh tekhnologiy: Doklad na II Mezhdunarodnom kongresse YuNESKO
“Obrazovanie i informatika” [State and prospects for the development of information and
telecommunication technologies: Report at the 2nd UNESCO Congress (Internation-
al) “Education and Information Science”]. *Informatsionnoe obshchestvo* [Information
Society] 5:3–22.
 37. Mizin, I. A., E. V. Kiselev, I. A. Sokolov, and S. Ya. Shorgin. 1996. Nekotorye
problemy sozdaniya edinoy informatsionno-telekommunikatsionnoy sistemy obshchena-
tsional'nogo masshtaba kak osnovy informatizatsii sfery obrazovaniya v Rossii [Some
problems of creating a unified information and telecommunications system of national
scale as a basis for informatization of the education sphere in Russia]. *Sistemy i Sredstva
Informatiki — Systems and Means of Informatics* 8:114–124.
 38. Kolin, K. K. 1996. Informatika v sisteme operezhayushchego obrazovaniya [Informatics
in the system of advanced education]. *Vestnik Rossiyskogo obshchestva informatiki
i vychislitel'noy tekhniki* [Bulletin of the Russian Society for Informatics and Computer
Engineering] 3:19–39. EDN: WLNXTV.
 39. Politika v oblasti obrazovaniya i novye informatsionnye tekhnologii: Natsional'nyy
doklad Rossii [Policy in education and new information technologies. National report of
Russia]. 1996. *UNESCO 2nd Congress (International) “Education and Informatics”
Proceedings*. Moscow. 34 p.
 40. Kolin, K. K. 1997. “Informatika-2000” — mezhdunarodnyy proekt dlya sistemy
obrazovaniya [“Informatics-2000” is an international project for education]. *Mezh-
dunarodnoe sotrudnichestvo* [International Cooperation] 1:17–18. EDN: WIXSSZ.

41. Kolin, K. K. 2022. Obrazovanie dlya informatsionnogo obshchestva: problemy i priority [Education for the information society: Problems and priorities]. *Informatsionnoe obshchestvo* [Information Society] 5:16–34. doi: 10.52605/16059921_2022_05_16. EDN: CIQJJS.

Received January 9, 2025

Accepted February 15, 2025

Contributors

Sokolov Igor A. (b. 1954) — Doctor of Science in technology, Academician of RAS, research advisor, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; isokolov@ipiran.ru

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

Zakharov Victor N. (b. 1948) — Doctor of Science in technology, associate professor, scientific secretary, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; vzakharov@ipiran.ru

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

МЕТОДЫ НОРМАЛЬНОЙ СУБОПТИМАЛЬНОЙ ФИЛЬТРАЦИИ В НАБЛЮДАЕМЫХ НЕЯВНЫХ ГАУССОВСКИХ СТОХАСТИЧЕСКИХ СИСТЕМАХ

*И. Н. Синицын*¹

Аннотация: Разработана теория нормальных субоптимальных фильтров (НСОФ) и модифицированных НСОФ (МНСОФ) для гауссовских непрерывных и дискретных неявных стохастических систем (СтС), приводимых к явным. Предполагается, что наблюдения не влияют на объект наблюдения и описываются явными стохастическими дифференциальными и разностными уравнениями. Дан обзор работ в области субоптимального синтеза НСОФ и МНСОФ первого типа для нелинейных и квазилинейных приведенных СтС с гладкими и разрывными неявными функциями. Разработаны алгоритмы синтеза НСОФ второго типа для приведенных систем на базе обобщенных фильтров Калмана (ОФК) и Калмана–Бьюси (ОФКБ). Нормальные субоптимальные фильтры второго типа, в отличие от НСОФ первого типа, не позволяют оценивать точность фильтрации заранее. Нормальные субоптимальные фильтры и МНСОФ могут найти применение в задачах быстрой обработки информации в технических и организационно-техничко-экономических системах с сосредоточенными параметрами невысокой размерности, когда постоянными времени при высших производных (разностях) можно пренебречь. Результаты допускают применение также в эредитарных системах, приводимых к дифференциальным системам (разностным). Предложены направления дальнейших исследований.

Ключевые слова: гауссовская стохастическая система (СтС); модифицированный НСОФ; неявная СтС; нормальные субоптимальные фильтры (НСОФ) первого и второго типа

DOI: 10.14357/08696527250102

EDN: GFRDWH

1 Введение

Как известно [1], практическое применение теории оптимальной нелинейной фильтрации для решения в реальном времени задач приближенной субоптимальной обработки информации в наблюдаемых явных дифференциальных СтС имеет смысл только в тех случаях, когда оценки можно вычислять в реальном времени по мере получения результатов наблюдений. Точная нелинейная теория дает оптимальные оценки в каждый момент времени по результатам наблюдений, полученных к этому моменту, без использования последующих результатов наблюдений. Если эти оценки не могут быть вычислены в тот же момент времени

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

или хотя бы с фиксированным приемлемым запаздыванием и их вычисление приходится откладывать на будущее, то нет никакого смысла отказываться от использования наблюдений, получаемых после момента времени, для оценивания состояния системы в момент времени. Поэтому для статистической обработки результатов после окончания наблюдений целесообразно применять известные из математической статистики методы постобработки информации.

Как известно [1], необходимость обработки результатов наблюдений в реальном времени непосредственно в процессе эксперимента привела к появлению ряда приближенных методов оптимальной нелинейной фильтрации, называемых обычно методами субоптимальной фильтрации. Приближенные методы субоптимальной фильтрации первого типа основаны на приближенном решении фильтрационных уравнений, а второго типа — на превращении формул для стохастических дифференциалов оптимальной оценки $\hat{X}_t$ и апостериорной ковариационной матрицы ошибки R_t в стохастические дифференциальные уравнения для $\hat{X}_t$ и R_t путем разложения нелинейных функций в ряды и отбрасывания остаточных членов.

Для приближенного решения уравнения для апостериорного одномерного вектора состояния X_t можно использовать методы, основанные на параметризации одномерного распределения стохастического процесса, определяемого стохастическим дифференциальным уравнением. Эти методы позволяют получить стохастические дифференциальные уравнения для параметров апостериорного распределения. Простейший из них — метод нормальной аппроксимации (МНА) апостериорного распределения. Другие методы основаны на использовании в качестве параметров апостериорных распределений моментов, семиинвариантов или коэффициентов ортогонального разложения апостериорной плотности вектора состояния системы. Исключительно важное практическое значение имеют квазилинейные субоптимальные фильтры (СОФ), получаемые с помощью методов эквивалентной линеаризации.

В [2–6] для неявных дифференциальных СтС, допускающих приведение к явным (т. е. не разрешенных относительно старших производных) в рамках СОФ первого типа разработаны методы синтеза НСОФ. В [7] для неявных СтС, приводимых к явным, со случайными параметрами, описываемыми многокомпонентными интегральными каноническими представлениями, разработаны методы синтеза и анализа точности НСОФ. Рассмотрим модифицированные для нормированных распределений методы субоптимального синтеза фильтров двух типов для неявных непрерывных и дискретных СтС, приводимых к явным. В разд. 2 рассматриваются дифференциальные и разностные неявные гауссовские СтС и их наблюдения. Нормальные СОФ первого типа рассматриваются в разд. 3. Раздел 4 посвящен МНСОФ первого типа. В разд. 5 рассматриваются квазилинейные НСОФ и МНСОФ. Нормальные СОФ второго типа представлены в разд. 6. В заключении приведены основные выводы и направления дальнейших исследований.

2 Приводимые неявные гауссовские стохастические системы и их наблюдения

Следуя [7], рассмотрим сначала векторную неявную дифференциальную СтС с нелинейностями, описываемыми гладкими детерминированными функциями размерности $(n^\Phi \times 1)$:

$$\Phi = \Phi(t, Z_t, \dot{Z}_t, \dots, Y_t^{(k)}, U_t) = 0, \\ Z(t_0) = Z_0, \dot{Z}(t_0) = \dot{Z}_0, \dots, Y^{(k)}(t_0) = Z_0^{(k)}. \quad (1)$$

При этом уравнение нелинейного формирующего фильтра для гауссовских стохастических возмущений U_t возьмем в виде, разрешенном относительно гауссовских белых шумов:

$$\dot{U}_t = a^U(t, U_t) + b^U(t, U_t) V_t^U, \quad U(t_0) = U_0. \quad (2)$$

Здесь $a^U(t, U_t)$ и $b^U(t, U_t)$ — $(n^Z \times 1)$ - и $(n^Z \times n^V)$ -мерные функции; V_t^U — гауссовский белый шум интенсивности ν_t . Уравнение (2) понимается в смысле Ито и имеет единственное среднееквадратичное решение [1].

Для гладких функций в (1), допускающих стохастические производные Ито до h -го порядка, выполним следующие преобразования. Будем дифференцировать по t левые части уравнений (1) по формуле Ито до тех пор, пока не появятся производные белого шума. В результате получим следующую систему неявных нелинейных дифференциальных уравнений:

$$\Phi = 0, \quad \dot{\Phi} = 0, \dots, \Phi^{(h)} = 0. \quad (3)$$

Далее введем составной вектор состояния $[\bar{Z}_t^T Z_t'^T]^T$, где

$$\bar{Z}_t = [Z_t^T \dot{Z}_t^T \dots Z_t^{(k-1)T}]^T$$

и вспомогательный вектор Z_t' определяется уравнениями (3) в виде конечного векторного уравнения:

$$\Psi(\bar{Z}_t, Z_t') = 0.$$

Таким образом, если неявные дифференциальные уравнения (1) в случае неявных гладких вектор-функций допускают гладкую замену переменных, то приведенные исходные неявные дифференциальные уравнения состоят из векторного стохастического дифференциального уравнения Ито для гауссовского белого шума $V = V_t^U$ и неявного векторного конечного уравнения (**теорема 2.1**):

$$\dot{\bar{Z}}_t = a^\Pi(\bar{Z}_t, t) + b^\Pi(\bar{Z}_t, t) V, \quad \Psi^\Pi(\bar{Z}_t, Z_t', t) = 0.$$

В задачах фильтрации и идентификации параметров в непрерывных СтС обычно принимается, что объектовая неявная дифференциальная СтС допускает приведение к явной СтС, измерительная система вполне наблюдаема и наблюдения не влияют на объект, причем уравнение наблюдения разрешено относительно вектора наблюдения. Введем новые обозначения для составных векторов состояния объекта X_t взамен $\bar{Z}_t$ ($X_t \equiv \bar{Z}_t$), вспомогательного вектора $\tilde{X}_t \equiv Z'_t$ и системы наблюдения Y_t . Тогда в качестве приведенных уравнений состояния неявного объекта и наблюдений будем рассматривать следующие уравнения (**теорема 2.2**):

$$\left. \begin{aligned} \dot{X}_t &= \varphi^\Pi(X_t, t) + \psi^\Pi(X_t, t) V, \quad \Psi^\Pi(X_t, \tilde{X}_t, t) = 0; \\ \dot{Y}_t &= \varphi_1(X_t, Y_t, t) + \psi_1(X_t, Y_t, t) V. \end{aligned} \right\} \quad (4)$$

Приведенные неявные дискретные СтС описываются разностными уравнениями вида (**теорема 2.3**)

$$\left. \begin{aligned} X_{k+1} &= \varphi_k^\Pi(X_k) + \psi_k^\Pi(X_k) V_k, \quad \Psi_k^\Pi(X_k, \tilde{X}_k) = 0; \\ Y_k &= \varphi_{1k}(X_k, Y_k) + \psi_{1k}(X_k, Y_k) V_k \quad (k = 1, 2, \dots). \end{aligned} \right\} \quad (5)$$

Замечание 2.1. Для негладких неявных функций в (1) используются методы линейной или полиномиальной регрессионной аппроксимации неявных функций [4, 7].

3 Нормальные субоптимальные фильтры первого типа

Применяя к приведенным уравнениям (4) теорию НСОФ [1], получим следующие результаты:

$$\begin{aligned} \dot{\hat{X}}_t &= f^\Pi(\hat{X}_t, R_t, t) + h^\Pi(\hat{X}_t, R_t, t) [\dot{Y}_t - f^{(1)}(\hat{X}_t, R_t, t)], \\ \Psi^\Pi(\hat{X}_t, \hat{\tilde{X}}_t) &= 0; \end{aligned} \quad (6)$$

$$\begin{aligned} \dot{R}_t &= \left\{ f^{\Pi(2)}(\hat{X}_t, R_t, t) - h^\Pi(\hat{X}_t, R_t, t) (\psi_1 \nu \psi_1^\top)(t) h^\Pi(\hat{X}_t, R_t, t)^\top \right\} + \\ &+ \sum_{r=1}^{n_y} \rho_r^\Pi(\hat{X}_t, R_t, t) [\dot{Y}_r - f_r^{\Pi(1)}(\hat{X}_t, R_t, t)]. \end{aligned} \quad (7)$$

Здесь введены обозначения:

$$f^{\Pi}(\hat{X}_t, R_t, t) = [(2\pi)^n |R_t|]^{-1/2} \int_{-\infty}^{\infty} \varphi^{\Pi}(x, t) \exp \left\{ -\frac{(x^T - \hat{X}_t^T) R_t^{-1} (x - \hat{X}_t)}{2} \right\} dx; \quad (8)$$

$$f^{(1)}(\hat{X}_t, R_t, t) = \left\{ f_r^{(1)}(\hat{X}_t, R_t, t) \right\} = [(2\pi)^{n_x} |R_t|]^{-1/2} \int_{-\infty}^{\infty} \varphi_1(x, t) \exp \left\{ -\frac{(x^T - \hat{X}_t^T) R_t^{-1} (x - \hat{X}_t)}{2} \right\} dx; \quad (9)$$

$$h^{\Pi}(\hat{X}_t, R_t, t) = \left\{ [(2\pi)^{n_x} |R_t|]^{-1/2} \int_{-\infty}^{\infty} \left[x \varphi_1(x, t)^T + (\psi^{\Pi} \nu \psi_1^T)(x, t) \right] \times \right. \\ \left. \times \exp \left\{ -\frac{(x^T - \hat{X}_t^T) R_t^{-1} (x - \hat{X}_t)}{2} \right\} dx - \right. \\ \left. - \hat{X}_t f^{\Pi(1)}(\hat{X}_t, Y_t, R_t, t)^T \right\} \left((\psi_1 \nu \psi_1^T) \right)^{-1}(t); \quad (10)$$

$$f^{\Pi(2)}(\hat{X}_t, R_t, t) = [(2\pi)^{n_x} |R_t|]^{-1/2} \int_{-\infty}^{\infty} \left\{ (x - \hat{X}_t) \varphi^{\Pi}(x, t)^T + \right. \\ \left. + \varphi^{\Pi}(x, t) (x^T - \hat{X}_t^T) + (\psi^{\Pi} \nu \psi_1^T)(x, t) \right\} \times \\ \times \exp \left\{ -\frac{(x^T - \hat{X}_t^T) R_t^{-1} (x - \hat{X}_t)}{2} \right\} dx; \quad (11)$$

$$\rho_r^{\Pi}(\hat{X}_t, R_t, t) = [(2\pi)^{n_x} |R_t|]^{-1/2} \int_{-\infty}^{\infty} \left\{ (x - \hat{X}_t) (x^T - \hat{X}_t^T) a_r(x, t) + \right. \\ \left. + (x - \hat{X}_t) b_r^{\Pi}(x, t)^T (x^T - \hat{X}_t^T) + b_r^{\Pi}(x, t) (x^T - \hat{X}_t^T) \right\} \times \\ \times \exp \left\{ -\frac{(x^T - \hat{X}_t^T) R_t^{-1} (x - \hat{X}_t)}{2} \right\} dx \quad (r = \overline{1, n_y}), \quad (12)$$

где a_r^Π — r -й элемент матрицы-строки $(\varphi_1^\Pi - \hat{\varphi}_1^\Pi)((\psi_1 \nu \psi_1^\Pi)^{-1})$; b_{kr}^Π — элемент k -й строки и r -го столбца матрицы $(\psi^\Pi \nu \psi_1^\Pi)((\psi_1 \nu \psi_1^\Pi)^{-1})$; b_r — r -й столбец матрицы $(\psi^\Pi \nu \psi_1^\Pi)((\psi_1 \nu \psi_1^\Pi)^{-1})$; $b_r^\Pi = [b_{1r}^\Pi \dots b_{pr}^\Pi]^\Pi$ ($r = \overline{1, n_1}$).

Число уравнений МНА одномерного апостериорного распределения определяется по формуле

$$Q_{\text{МНА}} = n_x + \frac{n_x(n_x + 1)}{2} = \frac{n_x(n_x + 3)}{2}. \quad (13)$$

За начальные значения $\hat{X}_t$ и R_t при интегрировании уравнений (6) и (7) следует принять условные математическое ожидание и ковариационную матрицу величины X_0 относительно Y_0 :

$$\hat{X}_0 = M[X_0 | Y_0]; \quad R_0 = M[(X_0 - \hat{X}_0)(X_0^\Pi - \hat{X}_0^\Pi) | Y_0]. \quad (14)$$

Если нет информации об условном распределении X_0 относительно Y_0 , то начальные условия можно взять в виде $\hat{X}_0 = M X_0$ и $R_0 = M(X_0 - M X_0)(X_0^\Pi - M X_0^\Pi)$. Если же и об этих величинах нет никакой информации, то начальные значения $\hat{X}_t$ и R_t приходится задавать произвольно.

Теорема 3.1. Пусть в условиях теоремы 2.1 приведенные нелинейные уравнения гауссовской СтС (4) допускают применение МНА, а коэффициенты (8)–(12) конечны. Тогда НСОФ определяется уравнениями (6), (7), при этом число уравнений вычисляется по формуле (13) при начальных условиях (14).

Замечание 3.1. Из формулы (12) видно, что если функция φ_1 линейна относительно X_t и функция ψ не зависит от X_t , то при нормальной аппроксимации все матрицы ρ_r равны нулю вследствие чего уравнения (7) не содержат $\dot{Y}_t$.

Замечание 3.2. Полагая в уравнении (7) $\dot{R}_t = 0$, получим уравнения для стационарных НСОФ.

Аналогично получаются дискретные уравнения НСОФ в случае приведенных разностных уравнений (5) (**теорема 3.2**):

$$\begin{aligned} \hat{X}_{k+1} &= f_k^\Pi(\hat{X}_{k+1|k}, R_{k+1|k}) + \\ &+ h_k^\Pi(\hat{X}_{k+1|k}, R_{k+1|k}) \left[Y_{k+1} - f_k^{(1)}(\hat{X}_{k+1|k}, R_{k+1|k}) \right], \quad \Psi_k^\Pi(\hat{X}_k \hat{X}_k) = 0; \\ R_{k+1} &= \left\{ f_k^{\Pi(2)}(\hat{X}_{k+1|k}, R_{k+1|k}) - \right. \\ &- h_k^\Pi(\hat{X}_{k+1|k}, R_{k+1|k}) (\psi_{1,k} \nu \psi_{1,k}^\Pi) h_k^\Pi(\hat{X}_{k+1|k}, R_{k+1|k})^\Pi \left. \right\} + \\ &+ \sum_{r=1}^{n_y} \rho_r^\Pi(\hat{X}_{k+1|k}, R_{k+1|k}) \left[Y_{r,k+1} - f_{r,k}^{\Pi(1)}(\hat{X}_{k+1|k}, R_{k+1|k}) \right]. \end{aligned}$$

Здесь введены следующие обозначения:

$$\begin{aligned}
 f_k^\Pi &= f_k^\Pi \left(\hat{X}_{k+1|k}, Y_k, R_{k+1|k} \right) = \mathbf{M} \left[\varphi_k^\Pi \right]; \\
 f_k^1 &= f_k^{(1)} \left(\hat{X}_{k+1|k}, Y_k, R_{k+1|k} \right) = \mathbf{M} \left[\varphi_{1,k} \right]; \\
 h_k^\Pi &= h_k^\Pi \left(\hat{X}_{k+1|k}, Y_k, R_{k+1|k} \right) = \mathbf{M} \left[x \varphi_{1k} (Y_k, x) + \psi_k \nu_k \psi_{1k}^\top (Y_k, x) \right]; \\
 f_k^{\Pi(2)} &= f_k^{\Pi(2)} \left(\hat{X}_{k+1|k}, Y_k, R_{k+1|k} \right) = \mathbf{M} \left[\left(x - \hat{X}_{k+1|k} \right) \varphi_k^\Pi (Y_k, x)^\top + \right. \\
 &\quad \left. + \varphi_k^\Pi (Y_k, x) \left(x^\top - \hat{X}_{k+1|k}^\top \right) + \psi_k^\Pi \nu_k \psi_{1,k}^\top (Y_k, x) \right]; \\
 \rho_r^\Pi &\left(\hat{X}_{k+1|k}, Y_k, R_{k+1|k} \right) = \mathbf{M} \left[\left(x - \hat{X}_{k+1|k} \right) \left(x^\top - \hat{X}_{k+1|k}^\top \right) a_r (Y_k, x) + \right. \\
 &\quad \left. + \left(x - \hat{X}_{k+1|k} \right) b_r^\Pi (Y_k, x)^\top \left(x^\top - \hat{X}_{k+1|k}^\top \right) + b_r^\Pi (Y_k, x) \left(x^\top - \hat{X}_{k+1|k}^\top \right) \right],
 \end{aligned}$$

где

$$\begin{aligned}
 \mathbf{M}[\bullet] &= \left[(2\pi)^{n_x} |R_{k+1|k}| \right]^{-1/2} \times \\
 &\times \int_{-\infty}^{\infty} (\bullet) \exp \left\{ -\frac{\left(x^\top - \hat{X}_{k+1|k}^\top \right) R_{k+1|k}^{-1} \left(x - \hat{X}_{k+1|k} \right)}{2} \right\} dx;
 \end{aligned}$$

a_r — r -й элемент матрицы строки $(\varphi_{1k}^\top - \hat{\varphi}_{1k}^\top)(\psi_{1k}\nu_k\psi_{1k}^\top)^{-1}$; $b_r^\Pi = [b_{1r}^\Pi \dots b_{n_x r}^\Pi]^\Pi$, b_{lr}^Π — элемент l -й строки и r -го столбца матрицы $\psi_k^\Pi \nu_k \psi_{1,k}^\top (\psi_{1k}\nu_k\psi_{1,k}^\top)^{-1}$.

В качестве начальных условий принимаются

$$\hat{X}_{1|1} = \hat{X}_1 = \mathbf{M} [X_1 | Y_1]; \quad R_{1|1} = R_1 = \mathbf{M} \left[(X_1 - X_1^0) (X_1 - X_1^0)^\top | Y_1 \right],$$

определяющие начальное нормальное распределение $\mathcal{N}(\hat{X}_1, R_1)$.

Число уравнений МНА одномерного распределения:

$$Q_{\text{МНА}} = \frac{n_x(n_x + 3)}{2}.$$

4 Модифицированные нормальные субоптимальные фильтры первого типа

В условиях теоремы 2.1, применяя теорию [1] МНСОФ к приведенным уравнениям (4), получим:

$$\left. \begin{aligned} \dot{m}_t &= \mu_t f_1^\Pi \left(\hat{X}_t, R_t, t \right) + h_1^\Pi \left(\hat{X}_t, R_t, t \right) \dot{Y}_t; \\ \dot{\Gamma}_t &= \mu_t f_2^\Pi \left(\hat{X}_t, R_t, t \right) + \mu_t \sum_{r=1}^{n_y} h_{2r}^\Pi \left(\hat{X}_t, R_t, t \right) \dot{Y}_{rt}; \\ \dot{\mu}_t &= \mu_t h_3 \left(\hat{X}_t, R_t, t \right) \dot{Y}_t. \end{aligned} \right\} \quad (15)$$

Здесь введены следующие обозначения:

$$\begin{aligned} f_1^\Pi &= f_1^\Pi \left(\hat{X}_t, R_t, t \right) = \\ &= \frac{1}{\sqrt{(2\pi)^{n_x} |R_t|}} \int_{-\infty}^{\infty} \varphi^\Pi(x, t) \exp \left\{ \frac{1}{2} \left(x^\mathrm{T} - \hat{X}_t^\mathrm{T} \right) R_t^{-1} \left(x - \hat{X}_t \right) \right\} dx; \end{aligned} \quad (16)$$

$$\begin{aligned} h_1^\Pi &= h_1^\Pi \left(\hat{X}_t, R_t, t \right) = \frac{1}{\sqrt{(2\pi)^{n_x} |R_t|}} \int_{-\infty}^{\infty} \left[\sigma_1^\Pi(x, t) + x \varphi_1(x, t) \right] \times \\ &\times \exp \left\{ \frac{1}{2} \left(x^\mathrm{T} - \hat{X}_t^\mathrm{T} \right) R_t^{-1} \left(x - \hat{X}_t \right) \right\} dx; \end{aligned} \quad (17)$$

$$\begin{aligned} f_2^\Pi &= f_2^\Pi \left(\hat{X}_t, R_t, t \right) = \\ &= \frac{1}{\sqrt{(2\pi)^{n_x} |R_t|}} \int_{-\infty}^{\infty} \left[x \varphi^\Pi(x, t)^\mathrm{T} + \varphi^\Pi(x, t) x^\mathrm{T} + \sigma_0^\Pi(x, R_t, t) \right] \times \\ &\times \exp \left\{ \frac{1}{2} \left(x^\mathrm{T} - \hat{X}_t^\mathrm{T} \right) R_t^{-1} \left(x - \hat{X}_t \right) \right\} dx \cdot \sigma_2(t)^{-1}; \end{aligned} \quad (18)$$

$$\begin{aligned} h_{2r}^\Pi &= h_{2r}^\Pi \left(\hat{X}_t, R_t, t \right) = \\ &= \frac{1}{\sqrt{(2\pi)^{n_x} |R_t|}} \int_{-\infty}^{\infty} \left[x x^\mathrm{T} a_r(x, t) + x b_r^\Pi(x, t)^\mathrm{T} + b_r^\Pi(x, t) x^\mathrm{T} \right] \times \\ &\times \exp \left\{ \frac{1}{2} \left(x^\mathrm{T} - \hat{X}_t^\mathrm{T} \right) R_t^{-1} \left(x - \hat{X}_t \right) \right\} dx \quad (r = \overline{1, n_x}); \end{aligned} \quad (19)$$

$$h_3 = h_3(\hat{X}_t, R_t, t) = \frac{1}{\sqrt{(2\pi)^{n_x} |R_t|}} \int_{-\infty}^{\infty} \varphi_1(x, t)^T \times \\ \times \exp \left\{ \frac{1}{2} (x^T - \hat{X}_t^T) R_t^{-1} (x - \hat{X}_t) \right\} dx \cdot \sigma_2(t)^{-1}; \quad (20)$$

$$\sigma_0^\Pi = \psi^\Pi \nu \psi^T; \quad \sigma_1^\Pi = \psi^\Pi \nu \psi_1^T; \quad \sigma_2 = \psi_1 \nu \psi_1^T, \quad (21)$$

где

$$\hat{X}_t = \mathbf{M} [X_t | Y_{t_0}^t], \quad \hat{X}_t = \frac{m_t}{\mu_t}; \quad (22)$$

$$R_t = \mathbf{M} \left[\left\{ X_t - \hat{X}_t \right\} X_t^T | Y_{t_0}^t \right], \quad R_t = \frac{\Gamma_t}{\mu_t} - \frac{m_t m_t^T}{\mu_t^2}; \quad (23)$$

$$\tilde{p}_t(x) \approx \tilde{p}_t^*(x) = \frac{\mu_t}{\sqrt{((2\pi)^n |R_t|)}} \exp \left\{ -\frac{1}{2} (x^T - \hat{X}_t^T) R_t^{-1} (x - \hat{X}_t) \right\}; \quad (24)$$

$$\tilde{g}_t(\lambda) \approx \tilde{g}_t^*(\lambda) = \mu_t \exp \left\{ i \lambda^T \hat{X}_t - \frac{1}{2} \lambda^T R_t \lambda \right\}; \quad (25)$$

$$m_r = m_{r_1, \dots, r_n} = \int_{-\infty}^{\infty} x_1^{r_1} \cdots x_n^{r_n} \tilde{p}_t(x) dx = \\ = \left[\frac{\partial^{|r|} \tilde{g}_t(\lambda)}{\partial (i \lambda_1)^{r_1} \cdots \partial (i \lambda_n)^{r_n}} \right]_{\lambda=0} = \mu \alpha_r; \quad (26)$$

$$\dot{\mu}_t = \mu_t \hat{\varphi}_1^T \left(\psi_1 \nu \psi_1^T \right)^{-1} (Y_t, t) \dot{Y}_t. \quad (27)$$

Уравнения (15)–(27) образуют замкнутую систему уравнений, приближенно определяющую m_t , Γ_t , μ_t , $\hat{X}_t$ и R_t . Начальными условиями служат

$$m(t_0) = m_0 = \hat{X}_0; \quad \Gamma(t_0) = \Gamma_0 = R_0 + \hat{X}_0 \hat{X}_0^T; \quad \mu(t_0) = 1, \quad (28)$$

где $\hat{X}_0$ и R_0 — условное математическое ожидание и ковариационная матрица вектора X_0 относительно Y_0 .

Уравнения (15)–(27) несколько проще, чем уравнения МНА при нормированном распределении, хотя и содержат на одно уравнение больше:

$$Q_{\text{ММНА}} = 1 + \frac{n_x(n_x + 3)}{2}. \quad (29)$$

Теорема 4.1. Пусть в условиях теоремы 2.1 приведенные нелинейные уравнения гауссовской СтС (4) допускают применение модифицированного МНА, а коэффициенты (16)–(20) конечны. Тогда МНСОФ определяется уравнениями (15) при условиях (28)–(29).

Аналогично для дискретной системы (5) имеем следующие уравнения **теоремы 4.2:**

$$\begin{aligned} m_{k+1} &= \mu_k f_{1k}^{\Pi} \left(\hat{X}_{k+1|k}, R_{k+1|k} \right) + \mu_k h_1^{\Pi} \left(\hat{X}_{k+1|k}, R_{k+1|k} \right); \\ \Gamma_{k+1} &= \mu_k f_2^{\Pi} \left(\hat{X}_{k+1|k} \right) + \mu_k \sum_{r=1}^{n_y} h_{2r}^{\Pi} \left(\hat{X}_{k+1|k}, R_{k+1|k} \right) Y_{r,k+1}; \\ \mu_{k+1} &= \mu_{k+1} + \mu_k h_3 \left(\hat{X}_{k+1|k}, R_{k+1|k} \right) (Y_{k+1} - Y_k). \end{aligned}$$

Здесь принято

$$\begin{aligned} f_{1k}^{\Pi} &= f_{1k}^{\Pi}(\hat{X}_{k+1|k}, R_{k+1|k}) = \mathbf{M} \left[\varphi_k^{\Pi}, x \right]; \\ h_{1k}^{\Pi} &= h_{1k}^{\Pi}(\hat{X}_{k+1|k}, R_{k+1|k}) = \mathbf{M} \left[\sigma_{1k}^{\Pi} + x \varphi_{1k}(x) \right]; \\ f_{2k}^{\Pi} &= f_{2k}^{\Pi}(\hat{X}_{k+1|k}, R_{k+1|k}) = \mathbf{M} \left[\left(x \varphi_k^{\Pi\Gamma} + \varphi_k^{\Pi} x^{\Gamma} + \sigma_{0k}^{\Pi} \right) (\sigma_{2k})^{-1}(x) \right]; \\ h_{2r} &= h_{2r}(\hat{X}_{k+1|k}, R_{k+1|k}) = \mathbf{M} \left[\left(x x^{\Gamma} + x b_r^{\Pi\Gamma} + b_r^{\Pi} x^{\Gamma} \right) (x) \right]; \\ h_3 &= h_3(\hat{X}_{k+1|k}, R_{k+1|k}) = \mathbf{M} \left[\varphi_{1k}(\sigma_{2k})^{-1}(x) \right]; \\ \sigma_{0k}^{\Pi} &= \psi_k^{\Pi} \nu_k \psi_k^{\Pi\Gamma}; \quad \sigma_{1k}^{\Pi} = \psi_k^{\Pi} \nu_k \psi_{1k}^{\Gamma}; \quad \sigma_{2k} = \psi_{1k} \nu_k \psi_{1k}^{\Gamma}. \end{aligned}$$

В качестве начальных условий следует принять

$$m_1 = m_{1|1} = \hat{X}_{1|1}; \quad \Gamma_1 = \Gamma_{1|1} = R_{1|1} + \hat{X}_{1|1} \hat{X}_{1|1}^{\Gamma}; \quad \mu_1 = 1,$$

определяющие начальное нормальное одномерное распределение.

5 Квазилинейные нормальные субоптимальные фильтры первого типа

Пусть функции φ^{Π} , ψ^{Π} , φ_1 и ψ_1 в (4) удовлетворяют условиям

$$\begin{aligned} \varphi^{\Pi} &= \varphi^{\Pi}(X_t, Y_t, t) = \varphi^{\Pi}(X_t, t); \quad \psi^{\Pi}(X_t, Y_t, t) V_0 = \psi^{\Pi}(t) V_1; \\ \varphi_1 &= \varphi_1(X_t, Y_t, t) = \varphi_1(X_t, t); \quad \psi_1(X_t, t) V_0 = V_2. \end{aligned}$$

Здесь V_1 и V_2 — независимые винеровские процессы размерности $n_{V_1} = n_x$ и $n_{V_2} = n_y$. Тогда будем иметь

$$\dot{X}_t = \varphi^\Pi(X_t, t) + \psi^\Pi(t)V_1; \quad Z_t = \dot{Y}_t = \varphi_1(X_t, t) + V_2. \quad (30)$$

Заменим уравнения (30) статистически линеаризованной системой, нелинейной относительно математических ожиданий m_t^x и m_t^z и линейной относительно центрированных составляющих $X_t^0 = X_t - m_t^x$ и $\hat{X}_t^0 = \hat{X}_t - \hat{m}_t^x$:

$$\begin{aligned} \dot{m}_t^x &= \varphi_{00}^\Pi(m_t^x, K_t^x, t); \\ m_t^z &= \varphi_{10}(m_t^x, K_t^x, t); \end{aligned} \quad (31)$$

$$\begin{aligned} \dot{X}_t^0 &= \varphi_{01}^\Pi(m_t^x, K_t^x, t) X_t^0 + \psi^\Pi(t)V_1; \\ Z_t^0 &= \varphi_{11}(m_t^x, K_t^x, t) X_t^0 + V_2, \end{aligned} \quad (32)$$

где $\varphi_{00}^\Pi = \varphi_{00}^\Pi(m_t^x, K_t^x, t)$, $\varphi_{10} = \varphi_{10}(m_t^x, K_t^x, t)$, $\varphi_{01}^\Pi(m_t^x, K_t^x, t) = \partial\varphi_0^\Pi/\partial m_t^x$ и $\varphi_{11}(m_t^x, K_t^x, t) = \partial\varphi_{10}/\partial m_t^x$ — коэффициенты статистической линеаризации нелинейных функций φ^Π и φ_1 , вычисляемые для нормального распределения $\mathcal{N}(m_t^x, K_t^x)$. При этом ковариационная матрица K_t^x будет определяться уравнением

$$\dot{K}_t^x = \varphi_{11}(m_t^x, K_t^x, t) K_t^x + K_t^x \varphi_{11}^\Pi(m_t^x, K_t^x, t)^\top + \psi^\Pi(t)\nu_1(t)\psi^\Pi(t)^\top.$$

Применяя уравнения линейного фильтра Калмана–Бьюси [1], получим искомые уравнения квазилинейного фильтра на основе метода статистической линеаризации (МСЛ) (**теорема 5.1**):

$$\begin{aligned} \dot{\hat{X}}_t &= \varphi_{00}^\Pi(m_t^x, K_t^x, t) - \varphi_{01}^\Pi(m_t^x, K_t^x, t) m_t^x + \varphi_{01}^\Pi(m_t^x, K_t^x, t) \hat{X}_t + \\ &+ R_t \varphi_{11}(m_t^x, K_t^x, t)^\top \nu_2(t)^{-1} [Z_t - \varphi_{11}(m_t^x, K_t^x, t) \hat{X}_t - \\ &- \varphi_{10}(m_t^x, K_t^x, t) + \varphi_{11}(m_t^x, K_t^x, t) m_t^x], \end{aligned}$$

где $\hat{X}_0 = \mathbf{M}X_0$;

$$\begin{aligned} \dot{R}_t &= \varphi_{01}^\Pi(m_t^x, K_t^x, t) R_t + R_t \varphi_{01}^\Pi(m_t^x, K_t^x, t)^\top - \\ &- R_t \varphi_{11}(m_t^x, K_t^x, t)^\top \nu_2(t)^{-1} \varphi_{11}(m_t^x, K_t^x, t) R_t + \\ &+ \psi^\Pi(t)\nu_1(t)\psi^\Pi(t)^\top, \quad R_0 = \mathbf{M} \left[(X_0 - \hat{X}_0) (X_0 - \hat{X}_0)^\top \right]. \end{aligned}$$

Входящие сюда m_t^x и K_t^x определяются из уравнений (31) и (32).

Для приводимой нелинейной гауссовской дискретной СтС вида

$$X_{k+1} = \varphi_k^\Pi(X_k) + \psi_k^\Pi V_{1k}; \quad Y_k = \varphi_{1k}(X_k) + V_{2k}$$

после статистической линеаризации нелинейных функций имеем:

$$\varphi_k^\Pi \approx \varphi_{0,k}^\Pi + \varphi_{1,k}^\Pi X_k^0; \quad \varphi_{1k} \approx \varphi_{10,k} + \varphi_{11,k} X_k^0.$$

Используя уравнения дискретного фильтра Калмана [1] при $a_{0k} = \varphi_{0,k}^\Pi$, $b_{0k} = \varphi_{10,k}$, $a_{1,k} = \varphi_{1,k}$ и $b_{1,k} = \varphi_{11,k}$, получаем искомые уравнения квазилинейного фильтра (**теорема 5.2**):

$$\begin{aligned} \hat{X}_{k+1} &= \hat{X}_{k+1|k} + \beta_{k+1} \left(Y_{k+1} - \varphi_{11,k} \hat{X}_{k+1|k} \right); \\ \hat{X}_{k+1|k} &= \varphi_{1,k} \hat{X}_k + \varphi_{0,k}^\Pi; \\ R_{k+1|k} &= \varphi_{1,k} R_k \varphi_{1,k}^T + \psi_k^\Pi \nu_{1k} \psi_k^{\Pi T}; \\ \beta_{k+1} &= R_{k+1|k} \varphi_{11,k+1}^T \left(\varphi_{11,k+1} R_{k+1|k} \varphi_{11,k+1}^T + \nu_{2,k+1} \right)^{-1}; \\ R_{k+1} &= R_{k+1|k} - \beta_{k+1} \varphi_{11,k+1} R_{k+1|k} = R_{k+1|k} - \\ &\quad - R_{k+1|k} \varphi_{11,k+1}^T (\varphi_{11,k+1} R_{k+1|k} \varphi_{11,k+1}^T + \nu_{2,k+1})^{-1} \varphi_{11,k+1} R_{k+1|k}. \end{aligned}$$

Замечание 5.1. Квазилинейные МНСОФ содержат на одно уравнение больше, чем НСОФ. При этом вычисление коэффициентов МНСОФ упрощается вследствие зависимости σ_0^Π , σ_1^Π , σ_2^Π , a_r^Π и b_r^Π только от времени. Аналогично рассматривается случай стационарных НСОФ.

Замечание 5.2. Теоремы 5.1 и 5.2 можно применять к приведенным уравнениям в случае параметрических небелых шумов, определяемых уравнениями формирующего фильтра, и билинейностей на основе эквивалентной линеаризации, в частности МСЛ [1].

Замечание 5.3. Вследствие нелинейности исходных моделей инструментальные переменные ковариационные матрицы R не будут в общем случае ковариационными матрицами ошибки фильтрации. Однако коэффициенты статистической линеаризации и инструментальные переменные не зависят от результатов наблюдений и могут быть определены до получения результатов наблюдений. Тогда линеаризованные уравнения могут быть проинтегрированы в реальном времени. При этом возможна априорная оценка точности квазилинейного НСОФ.

6 Субоптимальные фильтры второго типа

Для явных гладких дифференциальных СтС среди СОФ второго типа наиболее распространен ОФКБ и фильтры второго порядка [1]. В основе ОФКБ

лежит принцип линеаризации уравнений в окрестности неизвестной оценки $\hat{X}_t$. При этом придется линеаризировать только регулярные части уравнений (4), т. е. функции φ^Π и φ_1 относительно x , а функцию ψ^Π принять равной $\psi^\Pi(\hat{X}_t, \hat{\hat{X}}_t, t)$. Применяя теорию ОФКБ [1] к линеаризованным приведенным уравнениям, придем к следующим результатам.

Теорема 6.1. *Если в условиях теоремы 2.1 приведенные уравнения нелинейной гауссовской дифференциальной СтС (4) допускают линеаризацию в окрестности неизвестной оценки $\hat{X}_t$ и матрица $\sigma_1 = (\psi_1 \nu \psi_1^T)$ невырождена, то уравнения*

$$\left. \begin{aligned} \dot{\hat{X}}_t &= \varphi^\Pi(\hat{X}_t, t) + h^\Pi(\hat{X}_t, R_t, t) \left[\dot{Y}_t - \varphi_1(\hat{X}_t, t) \right], \quad \Psi^\Pi(\hat{X}_t, \hat{\hat{X}}_t) = 0; \\ \dot{R}_t &= \left\{ \varphi_x^\Pi(\hat{X}_t, t)^T R_t + R_t \varphi_x^\Pi(\hat{X}_t, R_t, t) + ((\psi_1 \nu \psi_1^T))(\hat{X}_t, t) - \right. \\ &\quad \left. - h^\Pi(\hat{X}_t, R_t, t) (\psi_1 \nu \psi_1^T)(t) h^\Pi(\hat{X}_t, R_t, t)^T \right\} dt, \end{aligned} \right\} \quad (33)$$

где введены обозначения

$$\varphi_x^\Pi = \frac{\partial}{\partial X_t} \varphi^\Pi(X_t, t)^T; \quad \varphi_{1x} = \frac{\partial}{\partial X_t} \varphi_1(X_t, t)^T;$$

$$h^\Pi(\hat{X}_t, R_t, t) = \left\{ R_t \varphi_{1x}(\hat{X}_t, t) + ((\psi^\Pi \nu \psi_1^T))(\hat{X}_t, t) \right\} (\psi_1 \nu \psi_1^T)^{-1}(Y_t, t),$$

определяют субоптимальный ОФКБ.

Уравнения (33) должны интегрироваться совместно при начальных условиях

$$\hat{X}(t_0) = \hat{X}_0 = M[X_0 | Y_0]; \quad R(t_0) = R_0 = M \left[(X_0 - \hat{X}_0) (X_0^T - \hat{X}_0^T) | Y_0 \right].$$

При этом ОФКБ имеет порядок $Q_{\text{ОКБ}} = Q_{\text{МНА}} = n(n+3)/2$, тот же, что и приближенно субоптимальный фильтр МНА апостериорного распределения (см. разд. 3).

Дискретный ОФК для гладкой приведенной системы (5) определяется следующей теоремой.

Теорема 6.2. *Если в условиях теоремы 2.3 уравнения гауссовской нелинейной дискретной СтС (5) допускают линеаризацию в окрестности неизвестной оценки $\hat{X}_k$ и матрица $\sigma_{1,k} = \psi_{1,k} \nu_k \psi_{1,k}^T$ невырождена, то уравнения дискретного ОФК имеют следующий вид:*

$$\hat{X}_{k+1} = \varphi_k^\Pi(\hat{X}_k) + \beta_{k+1} \left[Y_{k+1} - \varphi_{1,k+1}(\hat{X}_{k+1|k}) \right]; \quad \Psi_k(\hat{X}_k, \hat{\hat{X}}_k) = 0;$$

$$\begin{aligned}\beta_{k+1} &= \left[R_{k+1|k} \left(\frac{\partial \varphi_1^\Pi}{\partial X} \right)_{X=\hat{X}_{k+1|k}} + \psi_k^\Pi \nu_k \psi_{1,k}^\top \right] \left(\psi_{1,k} \nu_k \psi_{1,k}^\top \right)^{-1}; \\ R_{k+1} &= R_{k+1|k} - \beta_{k+1} \left(\frac{\partial \varphi_1^\Pi}{\partial X} \right)_{X=\hat{X}_{k+1|k}} R_{k+1|k}; \\ R_{k+1|k} &= \left(\frac{\partial \varphi^\Pi}{\partial X} \right)_{X=\hat{X}_k} R_k \left(\frac{\partial \varphi^\Pi}{\partial X} \right)_{X=\hat{X}_k}^\top + \psi_k^\Pi \left(\hat{X}_k \right) \nu_k \psi_k^\Pi \left(\hat{X}_k \right)^\top.\end{aligned}$$

Замечание 6.1. В случае негладких регулярных функций в уравнениях (5) используется МСЛ, а соответствующие результаты содержатся в разд. 5.

Замечание 6.2. Матрица R в уравнениях СОФ может и не быть близкой к апостериорной ковариационной матрице ошибки фильтрации. Однако это не дает возможности оценивать точность СОФ второго типа заранее. Вычислять апостериорную ковариационную матрицу ошибки R можно каждый раз только в процессе фильтрации.

7 Заключение

Для непрерывных и дискретных гауссовских СтС, приведенных к явным, разработана теория НСОФ и МНСОФ по среднеквадратичному критерию двух типов. Субоптимальные фильтры второго типа, в отличие от фильтров первого типа, предназначены для решения задач реального времени, но не позволяют оценивать точность фильтрации заранее. В качестве иллюстрации рассмотрены приведенные квазилинейные фильтры. Такие НСОФ могут найти применение в задачах быстрой обработки информации в технических и организационно-технично-экономических системах невысокой размерности [8–11], когда постоянными времени при высших производных (разностях) можно пренебречь. Результаты также допускают применение в эредитарных системах, приводимых к дифференциальным (разностным) системам [12–16].

Первые результаты по НСОФ и МНСОФ по среднеквадратичному критерию первого типа для частных случаев негауссовских неявных СтС, приводимых к дифференциальным, получены в [3]. Важное значение имеет развитие НСОФ и МНСОФ для типовых негауссовских неявных СтС невысокой размерности.

Нормальные СОФ и МНСОФ первого типа можно рассматривать как первое приближенное решение уравнений для одномерного апостериорного распределения (плотности или характеристической функции) теории нелинейной фильтрации [1]. Дальнейшим развитием СОФ представляются методы параметризации одномерного апостериорного распределения посредством методов моментов, ортогональных разложений и др., а также методы канонических представлений случайных функций [17].

Представляет интерес также развитие теории НСОФ по байесовым критериям для приводимых неявных СтС.

Литература

1. *Синицын И. Н.* Фильтры Калмана и Пугачёва. — 2-е изд. — М.: Логос, 2007. 776 с.
2. *Синицын И. Н.* Параметрическое аналитическое моделирование процессов в стохастических системах, не разрешенных относительно производных // Системы и средства информатики, 2017. Т. 27. № 1. С. 20–45. doi: 10.14357/08696527170102. EDN: YODCZL.
3. *Синицын И. Н.* Нормальные субоптимальные фильтры для дифференциальных стохастических систем, не разрешенных относительно производных // Информатика и её применения, 2021. Т. 15. Вып. 1. С. 3–10. doi: 10.14357/19922264210101. EDN: UPEHRI.
4. *Sinityn I. N.* Analytical modeling and estimation of normal processes defined by stochastic equations with unsolved derivatives // J. Mathematics Statistics Research, 2021. Vol. 3. Iss. 1. Art. 139. 7 p. doi: 10.36266/JMSR/139.
5. *Sinityn I. N.* Theory of control stochastic systems with unsolved derivatives // Automation and control. Theories and applications / Ed. E. Dadios. — IFSA Publishing S. L., 2021. P. 59–117. doi: 10.5772/intechopen.100448.
6. *Синицын И. Н.* Аналитическое моделирование и оценивание нестационарных нормальных процессов в стохастических системах, не разрешенных относительно производных // Системы и средства информатики, 2022. Т. 32. № 2. С. 58–71. doi: 10.14357/08696527220206. EDN: YMGERJ.
7. *Синицын И. Н.* Субоптимальная фильтрация в стохастических системах, не разрешенных относительно производных, со случайными параметрами // Информатика и её применения, 2024. Т. 18. Вып. 1. С. 2–10. doi: 10.14357/19922264240101. EDN: KUWMKJ.
8. *Евланов Л. Г., Константинов В. М.* Системы со случайными параметрами. — М.: Наука, 1976. 568 с.
9. Справочник по теории автоматического управления / Под ред. А. А. Красовского. — М.: Наука, 1987. 712 с.
10. *Александровская Л. Н., Аронов И. З., Круглов В. И. и др.* Безопасность и надежность технических систем. — М.: Университетская книга, Логос, 2008. 375 с.
11. *Синицын И. Н., Шаламов А. С.* Лекции по теории систем интегрированной логической поддержки. — 2-е изд. — М.: ТОРУС ПРЕСС, 2019. 1072 с. EDN: HVAZTX.
12. *Колмановский В. Б., Носов В. Р.* Устойчивость и периодические режимы регулируемых систем с последействием. — М.: Наука, 1981. 448 с.
13. *Азбелев Н. В., Максимов В. П., Рахматулина Л. Ф.* Введение в теорию функционально-дифференциальных уравнений. — М.: Наука, 1991. 277 с.
14. *Босов А. В.* Исследование робастности численных аппроксимаций фильтра Вон-эма // Информатика и её применения, 2023. Т. 17. Вып. 2. С. 41–49. doi: 10.14357/19922264230206. EDN: BGILKR.

15. Босов А. В. Оптимальная фильтрация состояния нелинейной динамической системы по наблюдениям со случайными запаздываниями // Информатика и её применения, 2023. Т. 17. Вып. 3. С. 8–17. doi: 10.14357/19922264230302. EDN: CFVYJM.
16. Босов А. В. Фильтрация состояния нелинейной динамической системы по наблюдениям со случайными запаздываниями // Автоматика и телемеханика, 2023. № 6. С. 49–66. doi: 10.31857/S000523102306003X. EDN: CSDMWT.
17. Синицын И. Н. Канонические представления случайных функций. Теория и применения. — 2-е изд. — М.: ТОРУС ПРЕСС, 2023. 816 с. doi: 10.30826/94588-308-6. EDN: XHBITA.

Поступила в редакцию 24.09.2024

Принята к публикации 15.02.2025

NORMAL SUBOPTIMAL FILTERING METHODS IN IMPLICIT OBSERVABLE GAUSSIAN STOCHASTIC SYSTEMS

I. N. Sinitsyn^{1,2}

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

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

Abstract: The article is dedicated to the theory of normal suboptimal filters (NSOF) and modified NSOF (MNSOF) for Gaussian continuous and discrete implicit stochastic systems (StS) reducible to explicit. It is supposed that observations do not influence the observable object and are described by explicit stochastic differential and difference equations. A short survey in the field of suboptimal NSOF (MNSOF) synthesis is given. Basic algorithms of NSOF (MNSOF) for the first type synthesis for nonlinear and quasi-linear reducible StS with smooth and nonsmooth implicit nonlinearities are described. The theory of NSOF of the second type for reducible implicit StS is developed on the basis of generalized Kalman–Bucy and Kalman filters. Such NSOF unlike NSOF (MNSOF) of the first type do not permit to estimate accuracy of filtering beforehand applications: quick (or real-time) information processing in technical or organization-technical-economical systems is described by small dimension equations when it is possible to neglect time constants at high derivatives (differences). The results are also applicable to implicit hereditary StS reducible to explicit differential (discrete) StS. Directions for future research are formulated.

Keywords: Gaussian stochastic system (StS); implicit StS; modified NSOF (MNSOF); normal suboptimal filter (NSOF) of the first and second types

DOI: 10.14357/08696527250102

EDN: GFRDWH

References

1. Sinitsyn, I. N. 2007. *Fil'try Kalmana i Pugacheva* [Kalman and Pugachev filters]. 2nd ed. Moscow: Logos. 776 p.
2. Sinitsyn, I. N. 2017. Parametricheskoe analiticheskoe modelirovanie protsessov v stokhasticheskikh sistemakh, ne razreshennykh otnositel'no proizvodnykh [Parametric analytical modeling of wide band processes in stochastic systems with unsolved derivatives]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 27(1):20–45. doi: 10.14357/08696527170102. EDN: YODCZL.
3. Sinitsyn, I. N. 2021. Normal'nye suboptimal'nye fil'try dlya differentsial'nykh stokhasticheskikh sistem, ne razreshennykh otnositel'no proizvodnykh [Normal suboptimal filtering for differential stochastic systems with unsolved derivatives]. *Informatika i ee Primeneniya — Inform. Appl.* 15(1):3–10. doi: 10.14357/19922264210101. EDN: UPEHRI.
4. Sinitsyn, I. N. 2021. Analytical modeling and estimation of normal processes defined by stochastic differential equations with unsolved derivatives. *J. Mathematics Statistics Research* 3(1):139. 7 p. doi: 10.36266/JMSR/139.
5. Sinitsyn, I. N. 2021. Theory of control stochastic systems with unsolved derivatives. *Automation and control. Theories and applications*. Ed. E. Dadios. IFSA Publishing S. L. 59–117. doi: 10.5772/intechopen.100448.
6. Sinitsyn, I. N. 2022. Analiticheskoe modelirovanie i otsenivanie nestatsionarnykh normal'nykh protsessov v stokhasticheskikh sistemakh, ne razreshennykh otnositel'no proizvodnykh [Analytical modeling and estimation of nonstationary normal processors with unsolved derivatives]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 32(2):58–71. doi: 10.14357/08696527220206. EDN: YMGERJ.
7. Sinitsyn, I. N. 2024. Suboptimal'naya fil'tratsiya v stokhasticheskikh sistemakh, ne razreshennykh otnositel'no proizvodnykh, so sluchaynymi parametrami [Suboptimal filtering in stochastic systems with random parameters and unsolved derivatives]. *Informatika i ee Primeneniya — Inform. Appl.* 18(1):2–10. doi: 10.14357/19922264240101. EDN: KUWMKJ.
8. Evlanov, A. G., and V. M. Konstantinov. 1976. *Sistemy so slozhnyimi parametrami* [Systems with random parameters]. Moscow: Nauka. 568 p.
9. Krasovskiy, A. A., ed. 1987. *Spravochnik po teorii avtomaticheskogo upravleniya* [Handbook for automatic control]. Moscow: Nauka. 712 p.
10. Aleksandrovskaya, L. N., I. Z. Aronov, V. I. Kruglov, et al. 2008. *Bezopasnost' i nadezhnost' tekhnicheskikh sistem* [Security and safety of technical systems]. Moscow: Universitetskaya kniga, Logos. 375 p.
11. Sinitsyn, I. N., and A. S. Shalamov. 2019. *Lektsii po teorii sistem integrirovannoy logisticheskoy podderzhki* [Lectures on theory of integrated logistic support systems]. 2nd ed. Moscow: TORUS PRESS. 1072 p. END: HVAZTX.
12. Kolmanovskiy, V. B., and V. R. Nosov. 1981. *Ustoychivost' i periodicheskie rezhimy reguliruemyykh sistem s posledeystviem* [Stability and periodic modes of regulated systems with aftereffects]. Moscow: Nauka. 448 p.
13. Azbelev, N. V., V. P. Maksimov, and L. F. Rakhmatulina. 1991. *Vvedenie v teoriyu funktsional'no-differentsial'nykh uravneniy* [Introduction to the theory of functional differential equations]. Moscow: Nauka. 277 p.

14. Bosov, A. V. 2023. Issledovanie robastnosti chislennykh approksimatsiy fil'tra Vonnama [Robustness investigation of the numerical approximation of the Wonham filter]. *Informatika i ee Primeneniya — Inform. Appl.* 17(2):41–49. doi: 10.14357/19922264230206. EDN: BGILKR.
15. Bosov, A. V. 2023. Optimal'naya fil'tratsiya sostoyaniya nelineynoy dinamicheskoy sistemy po nablyudeniym so sluchaynymi zapazdyvaniyami [Nonlinear dynamic system state optimal filtering by observations with random delays]. *Informatika i ee Primeneniya — Inform. Appl.* 17(3):8–17. doi: 10.14357/19922264230302. EDN: CFVYJM.
16. Bosov, A. V. 2023. Observation-based filtering of state of a nonlinear dynamical system with random delays. *Automat. Rem. Contr.* 84(6):594–605. doi: 10/1134/S0005117923060036. EDN: GVWEAB.
17. Sinitsyn, I. N. 2023. *Kanonicheskie predstavleniya sluchaynykh funktsiy. Teoriya i primeneniya* [Canonical expansion of random functions. Theory and application]. 2nd ed. Moscow: TORUS PRESS. 816 p. doi: 10.30826/94588-308-6. EDN: XHBITA.

Received September 24, 2024

Accepted February 15, 2025

Contributor

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

АНАЛИЗ СБОЕВ ПО КОСВЕННЫМ ПРИЗНАКАМ

*А. А. Грушо¹, Н. А. Грушо², М. И. Забезжайло³, Д. В. Смирнов⁴,
Е. Е. Тимонина⁵, С. Я. Шоргин⁶*

Аннотация: Информация о недоступности или неправильной работе сервиса для пользователя, которая получена в форме сообщения или жалобы в центр сбора такой информации, называется косвенным признаком сбоя. Такая информация генерируется пользователями сервисов, которые не ставят целью участие в изучении причин и устранении отказа в сервисе. Задачи анализа сбоев в больших распределенных информационных инфраструктурах по косвенным признакам актуальны, так как для мониторинга инфраструктуры привлекаются дополнительные объективные источники — пользователи сервисов. В статье построена формальная графовая модель информационной инфраструктуры (далее — системы), которая позволяет отражать сведения о появлениях сбоев с учетом локализации их в системе в зависимости от информативности косвенных признаков. Найденны условия выявления покрытия первопричины сбоя на различных уровнях иерархического описания системы, что означает идентификацию агрегированного объекта, в котором находится первопричина сбоя. Приведены примеры распространения сбоев и использования косвенных признаков для анализа нарушения работы системы при распространениях сбоев.

Ключевые слова: сбой; анализ сбоев по косвенным признакам; причинно-следственные связи

DOI: 10.14357/08696527250103

EDN: YWCEYW

1 Введение

Информация о недоступности или неправильной работе сервиса для пользователя, которая получена в форме сообщения или жалобы в центр сбора такой информации, называется косвенным признаком сбоя и может быть использована для идентификации области его локализации, а также диагностики причин возникновения [1]. Такая информация генерируется пользователями сервисов, которые не

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

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

³Федеральный исследовательский центр «Информатика и управление» Российской академии наук, m.zabezhailo@yandex.ru

⁴ПАО Сбербанк России, dvlsmirnov@sberbank.ru

⁵Федеральный исследовательский центр «Информатика и управление» Российской академии наук, eltimon@yandex.ru

⁶Федеральный исследовательский центр «Информатика и управление» Российской академии наук, sshorgin@ipiran.ru

ставят целью участие в изучении причин и устранении отказа в сервисе. Задачи анализа сбоев в больших распределенных информационных инфраструктурах по косвенным признакам актуальны [2–4], так как для мониторинга инфраструктуры привлекаются дополнительные объективные источники — пользователи сервисов. Это позволяет открыто собирать информацию о работоспособности всех элементов информационной инфраструктуры на больших территориях независимо от обслуживающих инфраструктуру сотрудников. С другой стороны, существует возможность фейков, а также проблем [5], связанных с тем, что организация сервисов с использованием информационной инфраструктуры создает систему зависимостей между организациями, предоставляющими сервисы и пользующимися сервисами, которую аналитически сложно описать. Поэтому, в частности, сбой трудно анализировать по единичным сообщениям о факте сбоя в предоставлении сервисов. Часто сбои происходят на прикладном уровне, хотя информационная система работает нормально. Примеры будут приведены далее. Для поиска первопричин сбоев даже в информационных инфраструктурах отдельных организаций информации о непредоставлении сервиса недостаточно.

В статье построена формальная графовая модель информационной инфраструктуры (далее — системы), которая позволяет отражать сведения о появлениях сбоев с учетом локализации их в системе в зависимости от информативности косвенных признаков.

Найдены условия выявления покрытия первопричины [6] сбоя на различных уровнях иерархического описания системы, что означает идентификацию агрегированного объекта, в котором находится первопричина сбоя.

Приведены примеры распространения сбоев и использования косвенных признаков для анализа нарушения работы системы при распространениях сбоев.

2 Модель информационной инфраструктуры и сбоев

Определим следующие конечные множества. Пусть P — множество базовых свойств системы; O — базовое множество объектов системы: если o — объект системы, то он состоит из свойств множества P . Объекты и свойства имеют индивидуальные идентификаторы. Объекты имеют смысл узлов и подсистем системы, объекты описываются свойствами узлов и подсистем. Состав свойств в объекте может изменяться.

Объекты базового множества O не пересекаются между собой. Если O — конечное множество базовых объектов, то допускаем следующее расширение этого множества O^* : объединение базовых объектов также считается объектом. Точно так же объединения базовых свойств P^* также считаем свойствами.

Множество свойств P_v из P будем называть *множеством видимых свойств*, т. е. таких свойств, которые однозначно идентифицируются в объекте, в котором они встречаются. Видимые свойства идентифицируются также в объектах, включающих объекты, содержащие видимое свойство.

Множество базовых объектов O образует множество вершин ориентированного графа G , дуги графа называются связями или зависимостями, т. е. возможностями одних объектов информационно воздействовать на другие. Новые свойства могут возникать в объекте или исчезать.

Существуют два механизма возникновения или исчезновения свойств в базовом объекте. Свойства возникают или исчезают в объекте случайно (с распределениями, зависящими от идентификатора объекта), и одни свойства могут порождать другие свойства даже в других объектах. Свойство p_1 в объекте o_1 может породить свойство p_2 в объекте o_2 тогда и только тогда, когда есть ориентированная связь из o_1 в o_2 и p_1 служит причиной p_2 , т. е. при наличии такой ориентированной связи p_1 порождает в o_2 свойство p_2 . Возможна ситуация, когда p_1 лежит в объекте o_1 , p_2 лежит в объекте o_2 , из o_1 и o_2 в графе G идут дуги в o_3 , а свойство $p_1 \cup p_2$ становится причиной свойства p_3 , которое появляется в объекте o_3 .

Отсюда возникает причинно-следственный граф C на множестве вершин P . Граф C также можно агрегировать следующим образом. Если в графе G присутствует дуга от o_1 к o_2 , то это значит, что существует объект o_3 , не совпадающий с o_1 и o_2 , который физически реализует связь объекта o_1 с объектом o_2 . Объект o_3 не будем отображать отдельной вершиной, но в случаях необходимости будем наносить его на дуги графа o_1 . Порождение свойства требует существования связи и наличия причинно-следственной связи.

Таким образом, граф G на множестве базовых объектов O и граф C на множестве базовых свойств P определяют возможные причинно-следственные связи в системе.

Граф G служит основой иерархии графов. Заменяем в графе G множество объектов φ на один объект $o(\varphi)$, представляющий собой их объединение. Если из объекта o в множество объектов φ вели несколько дуг, то заменяем их одной дугой (в случае необходимости с меткой, состоящей из множества имен объектов, в которые из o в φ входили дуги). Аналогично поступим с дугами, которые выходят из объектов множества φ . Создадим несколько агрегированных таким образом новых вершин и распространим на них описанные правила построения дуг. Получим новый агрегированный граф G^* , который однозначно определяется вершинами $o(\varphi)$ для различных непересекающихся φ и построенными между ними дугами.

Определим понятие *сервиса*. Пусть пользователь объекта o_1 хочет получить сервис, который может предоставить объект o_2 . Необходимое условие такой возможности — наличие связи от o_1 к o_2 и от o_2 к o_1 . Будем считать, что обеспечение таких связей выполняет объект o_3 . Для обращения к сервису o_2 пользователь создает в объекте o_1 свойство p_1 , которое становится причиной свойства p_2 в объекте o_2 . Тогда при наличии связи свойство p_1 порождает в o_2 свойство p_2 , которое можно назвать запросом на сервис. Свойство p_2 ввиду внутренней причинно-следственной связи создает в o_2 свойство p_3 , которое можно интерпретировать как ответ на запрос сервиса, и через связь o_2 к o_1 создает в силу

причинно-следственной связи свойство p_4 в o_1 . Свойство p_4 интерпретируется как выполнение запрошенного o_1 сервиса.

Сбой в информационной инфраструктуре в научной литературе определяется по-разному. Это связано с многообразием нарушений работы в различных технических системах, каждое из которых требует своего уточнения в терминологии [7]. В настоящей работе рассматривается ограниченное множество нарушений работы, поэтому выбран единый термин, который определен ниже.

Пусть $o \in O^*$ и $p \in o$. Тогда *сбой* (o, p) в объекте o — это исчезновение или существенная модификация p^{**} свойства p из o . Если модификация создала свойство p^{**} в o , то это свойство может перейти в o_1 вместо следствия p_1 причины p при наличии связи и причинно-следственной связи порождения причиной p свойства p_1 .

Сбой может быть случайным, тогда *первопричина* сбоя находится в o , или оказаться следствием некоторой причины, которая могла находиться в другом объекте.

Возможно, что причина случайно исчезла, тогда при наличии связи исчезнет следствие. Наоборот, возникновение сбоя p^{**} в одном объекте может породить сбой p^{**} в другом при наличии связи и отношения «причина–следствие». Для простоты будем считать, что если исходное свойство p_1 в объекте o_1 порождало свойство p_2 в объекте o_2 и p_1 модифицировалось в свойство p^{**} , то в объекте o_2 вместо p_2 появилось свойство p^{**} . При этом предполагаем, что объект связи o_3 не реагирует на передачу p^{**} , так как система связи, как правило, не зависит от передаваемой информации. Поэтому можно считать, что сбой системы передачи информации сводится только к одному виду сбоя — исчезновению способности к передаче информации [8].

Отметим, что в описанной модели возник механизм распространения сбоя. Добавим возможность появления очередей. Если в объекте o_2 вместо свойства предоставления сервиса p_2 возникло свойство p^{**} , то «зависают» обращения других объектов к сервису p_2 . Вследствие этого образовалась очередь на обслуживание сервисом p_2 . Это один из способов распространения сбоя.

3 Сбои и косвенные признаки сбоев

Пример 1. Использование модели инфраструктуры для описания порождения жалоб.

1. Если пользователь объекта o_1 генерирует в o_1 свойство p_1 для использования сервиса в объекте o_2 , то через объект o_3 , который предоставляет o_1 и o_2 услуги сети, порождается свойство p_2 в o_2 ввиду причинно-следственной связи, которое, в свою очередь, порождает результат p сервиса «ответ для o_1 » и передает его через o_3 для o_1 . Если в o_2 вместо свойства p_2 возникает свойство из p^{**} , то o_1 не получает p . Тогда пользователь o_1 генерирует жалобу $J(o_1)$ на отсутствие сервиса o_2 .

Построенную модель можно рассматривать как организацию базы данных для сбора и накопления информации о сбоях.

Модификация p_1 , не позволяющая отправить p_1 к o_1 в вершине (объекте) графа G или в вершине агрегированного графа G^* , означает появление сбоя, который не генерирует видимое свойство, т. е. означает исчезновение видимого свойства p_1 .

Пусть возникшее вместо свойства p_1 видимое свойство p^{**} в объекте o_1 означает невозможность сформировать операцию по обращению к сервису, предоставляемому объектом o_2 , т. е. o_1 не может обратиться к o_2 по имеющемуся каналу связи. Тогда свойство p^{**} не может возникнуть в o_2 как следствие сбоя в o_1 . Этот сигнал выделяет объект o_1 , в котором может присутствовать сбой (пользователь ошибся в управлении обращением к сервису, но связь по-прежнему возможна, иначе нельзя передать жалобу). Эта информация становится первым приближением в анализе сбоя. Далее возможны следующие варианты:

- (а) сбой произошел в реализации связи в o_3 с исполнителем сервиса o_2 (объект o_3 — это связь o_1 с o_2);
- (б) сбой произошел на приемном конце в объекте o_2 . Этот случай будет рассмотрен позже.

Не имея дополнительной информации, можно утверждать, что первопричина сбоя покрывается объектом $o_1 \cup o_2 \cup o_3$, хотя порождается одной жалобой.

2. Для оценки надежности жалобы потребуется дополнительная информация. Для анализа прежде всего необходимо получить информацию о работающих сервисах. Например, эту информацию можно открыто получить из предварительной оценки интенсивности их работы, т. е. для каждого сервиса x определим величину $t(x)$ как время, когда с вероятностью, достаточно близкой к 1, должен прийти запрос на использование x . При этом считаем, что множество сервисов конечно и такие оценки можно построить для всех регионов сбора данных. Тогда полученную жалобу на отсутствие сервиса от o_1 можно соотнести с работающими сервисами, если на них не пришли жалобы. Отдельно пришедшая жалоба скорее всего означает ошибку в управлении или сбой в компьютере пользователя объекта o_1 . Если жалоб на этот сервис от других источников нет, то, используя величину $t(x)$, приходим к выводу, что связь o_3 и объект o_2 работают исправно. Тогда объект o_1 объявляется результатом анализа причины сбоя. При этом видимое свойство сигнала сбоя может исчезнуть. Если одинаковых жалоб на сервис несколько от различных исходных адресов, то это, возможно, случаи (а) или (б). Каждая из возможных причин может быть проверена контрафактуальным [9] наблюдением следствий рассматриваемой причины или их отсутствием. Рабочее состояние функции связи отслеживается по отсутствию жалоб в промежутке времени $t(x)$ для зависимых наблюдаемых объектов.

Приведенный анализ носит вероятностный характер, так как нужный для анализа сервис может не использоваться или анализируемый собой может быть исправлен во время анализа. Однако увеличение числа жалоб на неработающий сервис служит подтверждением покрытия причины объектом $o_2 \cup o_3$. Результат анализа сбоя всегда представляет собой объект, покрывающий причину сбоя. Это объясняется тем, что более детальный анализ связан с получением информации о внутренней архитектуре покрывающих причину сбоя объектов, но эта информация, как правило, конфиденциальна и недоступна.

4 Использование косвенных данных при распространении сбоя

Распространение сбоя представляет наибольшую опасность в ущербе от сбоев. Приведем примеры механизмов распространения сбоев и возможностей их выявления с помощью жалоб.

Пример 2. Короткая цепочка распространения сбоя.

Пусть в примере 1 сбой произошел в объекте o_2 , но для ответа объекту o_1 объект o_2 должен был обратиться к объекту o_5 , который служит базой данных необходимой для ответа информации. Возможны два варианта дальнейшего развития ситуации.

В первом случае свойство из p^{**} , появившееся в o_2 , не может организовать запрос в базу данных объекта o_5 . В этом случае база данных не будет повреждена и продолжит работать для обслуживания запросов других объектов. Первопричина сбоя в этом случае покрывается объектом $o_2 \cup o_3$, что сразу следует из нескольких жалоб на недоступность сервиса в o_2 .

Во втором случае объект o_2 имеет возможность передать собой объекту o_5 , который уже не может обслуживать другие сервисы. Такая ситуация может возникнуть, например, если информация от o_2 блокирует ввод данных для обращения других пользователей. Возможная реакция будет выражаться в том, что другие объекты воспримут задержку как занятость входного интерфейса. Тогда часть системы будет блокирована ожиданием в очереди. При множестве жалоб на сервис o_2 и o_4 первопричина сбоя будет покрыта объектом $o_2 \cup o_4$.

Пример 3. Массовый вывод из строя сервисов в системе с помощью вредоносного кода.

Считается, что бот-сети предназначены для DDoS-атак на отдельные компьютеры или для распространения рекламы. Однако бот-сети можно использовать для нарушения работы множества различных сервисов. В простейшем случае бот может добавлять в пароли и контрольные коды дополнительные знаки, которые не позволят обращаться к любым сервисам, требующим аутентификации [10]. Этот факт выявляется потоком жалоб. При этом сохраняется работоспособность для других задач, т. е. чистый от бота объект o_5 получит доступ к сервису. Следовательно, первопричина кроется в зараженности других компьютеров.

Пример 4. Вирусная атака для распространения сбоя.

Предположим, что объект o_1 смог заразить объект o_2 вирусом, выполняющим следующие функции. Сервис o_2 может заражать полученным вирусом объекты, которые к нему обращаются за сервисом. В свою очередь, эти объекты могут заразить вирусом другие объекты, предоставляющие другие сервисы, т. е. возможно распространение вируса по объектам, предоставляющим сервисы [11]. Далее возможны вариации, когда очередное заражение приходит на зараженную машину. Например, происходит искажение пароля при обращении за сервисом. Таким образом возникает механизм постепенного поражения системы, причем при сохранении работоспособности каждого отдельного компьютера. Если присутствие вируса сохраняется при перезагрузке, то система оказывается полностью пораженной [12]. В этом случае поиск первопричины не позволит восстановить работоспособность системы.

Поиск первопричины не связан с множеством жалоб. Однако развитие процесса поражения с помощью жалоб может дать информацию о последовательности поражений [7].

Пример 5. Распространение сбоя с помощью очередей.

В работе [13] описан пример полного сбоя облачной системы на платформе Open Stack. В основе полного выхода из строя облака, когда все показатели выполнения функций помечены символом «ошибка», лежало изменение настройки одной управляющей программы, которая создала очередь в системе коммутации сообщений. Поскольку программа должна обслуживать все потоки, то очередь начала быстро разрастаться, создавая производные очереди на других узлах. Результат носил тотальный характер, причем поиски первопричины заняли много времени, несмотря на высокий уровень специалистов.

В рассмотренном примере жалобу формально мог бы послать любой пользователь, но она не смогла бы помочь в восстановлении системы. Однако тотальное отключение системы исключает существование жалоб, но порождает жалобы от внешних по отношению к системе объектов, которые пытаются обратиться к системе из других информационных пространств [14].

5 Заключение

Основная идея построенной модели состоит в возможности автоматического анализа поступающих сигналов и жалоб о сбоях сервисов. Разумеется, эта модель должна быть дополнена моделью нормализации жалоб и сигналов, т. е. должна существовать система приведения такой информации к единому языку для входной информации. В силу функциональной ограниченности обсуждаемых сервисов и архитектур такие решения уже существуют на практике в IBM [15], в Сбербанке [16] и в Яндексe [17].

Модель, представленная в разд. 2, может развиваться и масштабироваться. Кроме того, эта модель может быть распространена на более сложные и практически значимые случаи.

Также в статье рассмотрены варианты распространения сбоев. Самым опасным считается случай каскадного распространения сбоя. Этот случай просто выявляется с помощью потока жалоб примера 5, но выявление первопричины с помощью множества жалоб представляется маловероятным.

Литература

1. Грушо А. А., Забежайло М. И., Тимонина Е. Е. О каузальной репрезентативности обучающих выборок прецедентов в задачах диагностического типа // Информатика и её применения, 2020. Т. 14. Вып. 1. С. 80–86. doi: 10.14357/19922264200111. EDN: LAQCSMA.
2. Грушо А. А., Забежайло М. И., Смирнов Д. В., Тимонина Е. Е. Интеллектуальный анализ пополняемых коллекций Big Data в режиме процессно-реального времени // Информатика и её применения, 2021. Т. 15. Вып. 2. С. 36–43. doi: 10.14357/19922264210206. EDN: RTPKUR.
3. Грушо А. А., Грушо Н. А., Забежайло М. И., Писковский В. О., Тимонина Е. Е., Шоргин С. Я. Логика обмана в машинном обучении // Информатика и её применения, 2024. Т. 18. Вып. 1. С. 78–83. doi: 10.14357/19922264240111. EDN: SCDSHX.
4. Смирнов Д. В., Грушо А. А., Забежайло М. И. К задаче идентификации сбоев в информационно-технологической инфраструктуре путем мониторинга и анализа косвенных данных // Системы и средства информатики, 2024. Т. 34. № 3. С. 14–22. doi: 10.14357/08696527240302. EDN: QZYOGK.
5. Rinaldi S. M., Peerenboom J. P., Kelly T. K. Identifying, understanding, and analyzing critical infrastructure interdependencies // IEEE Contr. Syst. Mag., 2001. Vol. 21. No. 6. P. 11–25. doi: 10.1109/37.969131.
6. Грушо А. А., Грушо Н. А., Забежайло М. И., Кульченков В. В., Тимонина Е. Е. Выявление причинно-следственных связей при покрытии причин // Информатика и её применения, 2024. Т. 18. Вып. 2. С. 54–59. doi: 10.14357/19922264240208. EDN: MKXMZY.
7. Шубинский И. Б. Активная защита от отказов управляющих модульных вычислительных систем. — СПб.: Наука, 1993. 284 с.
8. Грушо А. А., Грушо Н. А., Забежайло М. И., Кульченков В. В., Тимонина Е. Е., Шоргин С. Я. Причинно-следственные связи в задачах классификации // Информатика и её применения, 2023. Т. 17. Вып. 1. С. 43–49. doi: 10.14357/19922264230106. EDN: DTQZPK.
9. Höfler M. Causal inference based on counterfactuals // BMC Med. Res. Methodol., 2005. Vol. 5. Art. 28. doi: 10.1186/1471-2288-5-28.
10. Грушо А. А., Грушо Н. А., Забежайло М. И., Тимонина Е. Е., Шоргин С. Я. Сложные причинно-следственные связи // Информатика и её применения, 2023. Т. 17. Вып. 2. С. 84–89. doi: 10.14357/19922264230212. EDN: TGXQIW.
11. Грушо А. А., Грушо Н. А., Забежайло М. И., Зацаринный А. А., Тимонина Е. Е., Шоргин С. Я. Анализ цепочек причинно-следственных связей // Информатика и её применения, 2022. Т. 16. Вып. 2. С. 68–74. doi: 10.14357/19922264220209. EDN: HPSLTU.

12. 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. Т. 10. С. 33–45. EDN: LHSORH.
13. 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. doi: 10.14357/08696527170203. EDN: YPJBAR.
14. Грушо А. А., Забейайло М. И., Смирнов Д. В., Тимонина Е. Е. Модель множества информационных пространств в задаче поиска инсайдера // Информатика и её применения, 2017. Т. 11. Вып. 4. С. 65–69. doi: 10.14357/19922264170408. EDN:
15. IBM Watson. <https://www.ibm.com/watson/services/natural-language-understanding>.
16. Сбер GigaChat. <https://developers.sber.ru/portal/products/gigachat>.
17. YandexGPT. <https://yandex.cloud/en/services/yandexgpt>.

Поступила в редакцию 09.01.2025

Принята к публикации 15.02.2025

ANALYSIS OF FAILURES BY INDIRECT FEATURES

*A. A. Grusho¹, N. A. Grusho¹, M. I. Zabezhalo¹, D. V. Smirnov², E. E. Timonina¹,
and S. Ya. Shorgin¹*

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

²Sberbank of Russia, 19 Vavilov Str., Moscow 117999, Russian Federation

Abstract: Information about the unavailability or incorrect operation of the service for the user, which is received in the form of a message or complaint to the center for collecting such information, is called an indirect feature of a failure. Such information is generated by users of services that are not to have the aim in participating in studying the causes and eliminating the failure in the service. The tasks of analyzing failures in large distributed information infrastructures by indirect features are relevant, since additional objective sources are involved in monitoring the infrastructure — users of services. The article builds a formal graph model of the information infrastructure (hereinafter, the system), which allows one to reflect information about the appearance of failures, taking into account their localization in the system, depending on the information content of indirect features. The conditions were found to detect the root cause of the failure at various levels of the hierarchical description of the system that means the identification of the aggregated object in which the root cause of the failure

is located. The examples of propagation of failures and the usage of indirect features to analyze system malfunction during fault propagation are given.

Keywords: failures; analysis of failures by indirect features; cause and effect

DOI: 10.14357/08696527250103

EDN: YWCEYW

References

1. Grusho, A. A., M. I. Zabezhalo, and E. E. Timonina. 2020. O kauzal'noy reprezentativnosti obuchayushchikh vyborok pretsedentov v zadachakh diagnosticheskogo tipa [On causal representativeness of training samples of precedents in diagnostic type tasks]. *Informatika i ee Primeneniya — Inform. Appl.* 14(1):80–86. doi: 10.14357/1992264200111. EDN: LAQCMA.
2. Grusho, A. A., M. I. Zabezhalo, D. V. Smirnov, and E. E. Timonina. 2021. Intel'lektual'nyy analiz popolnyaemykh kollektsey Big Data v rezhime protsessno-real'nogo vremeni [Intelligent analysis of Big Data extendible collections under the limits of process-real time]. *Informatika i ee Primeneniya — Inform. Appl.* 15(2):36–43. doi: 10.14357/1992264210206. EDN: RTPKUR.
3. Grusho, A. A., N. A. Grusho, M. I. Zabezhalo, V. O. Piskovski, E. E. Timonina, and S. Ya. Shorgin. 2024. Logika obmana v mashinnom obuchenii [Logic of deception in machine learning]. *Informatika i ee Primeneniya — Inform. Appl.* 18(1):78–83. doi: 10.14357/1992264240111. EDN: SCDSHX.
4. Smirnov, D. V., A. A. Grusho, and M. I. Zabezhalo. 2024. K zadache identifikatsii sboev v informatsionno-tekhnologicheskoy infrastrukture putem monitoringa i analiza kosvennykh dannykh [To the problem of identifying failures in the information technology infrastructure by monitoring and analyzing indirect data]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 34(3):14–22. doi: 10.14357/08696527240302. EDN: QZYOGK.
5. Rinaldi, S. M., J. P. Peerenboom, and T. K. Kelly. 2001. Identifying, understanding, and analyzing critical infrastructure interdependencies. *IEEE Contr. Syst. Mag.* 21(6):11–25. doi: 10.1109/37.969131.
6. Grusho, A. A., N. A. Grusho, M. I. Zabezhalo, V. V. Kulchenkov, and E. E. Timonina. 2024. Vyyavlenie prichinno-sledstvennykh svyazey pri pokrytii prichin [Identification of cause-and-effect relationships when covering causes]. *Informatika i ee Primeneniya — Inform. Appl.* 18(2):54–59. doi: 10.14357/1992264240208. EDN: MKXMZY.
7. Shubinskiy, I. B. 1993. *Aktivnaya zashchita ot otkazov upravlyayushchikh modul'nykh vychislitel'nykh sistem* [Active protection against failures of control modular computing systems]. Saint Petersburg: Nauka. 284 p.
8. Grusho, A. A., N. A. Grusho, M. I. Zabezhalo, V. V. Kulchenkov, E. E. Timonina, and S. Ya. Shorgin. 2023. Prichinno-sledstvennye svyazi v zadachakh klassifikatsii [Causal relationships in classification problems]. *Informatika i ee Primeneniya — Inform. Appl.* 17(1):43–49. doi: 10.14357/1992264230106. EDN: DTQZPK.
9. Höfler, M. 2005. Causal inference based on counterfactuals. *BMC Med. Res. Methodol.* 5:28. 12 p. doi: 10.1186/1471-2288-5-28.
10. Grusho, A. A., N. A. Grusho, M. I. Zabezhalo, E. E. Timonina, and S. Ya. Shorgin. 2023. Slozhnye prichinno-sledstvennye svyazi [Complex cause-and-effect relationships]. *Informatika i ee Primeneniya — Inform. Appl.* 17(2):84–89. doi: 10.14357/1992264230212. EDN: TGXQIW.

11. Grusho, A. A., N. A. Grusho, M. I. Zabezhailo, A. A. Zatsarinny, E. E. Timonina, and S. Ya. Shorgin. 2022. Analiz tsepochek prichinno-sledstvennykh svyazey [Cause-and-effect chain analysis]. *Informatika i ee Primeneniya — Inform. Appl.* 16(2):68–74. doi: 10.14357/19922264220209. EDN: HPSLTU.
12. Grusho, A. A., N. A. Grusho, and E. 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. EDN: LHSORH.
13. Grusho, A. A., M. I. Zabezhailo, A. A. Zatsarinny, A. V. Nikolaev, V. O. Piskovski, and E. E. Timonina. 2017. Klassifikatsiya oshibochnykh sostoyaniy v raspredelennykh vychislitel'nykh sistemakh i istochniki ikh vzniknoveniya [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. doi: 10.14357/08696527170203. EDN: YPJBP.
14. Grusho, A. A., M. I. Zabezhailo, D. V. Smirnov, and E. E. Timonina. 2017. Model' mnozhestva informatsionnykh prostranstv v zadache poiska insaydera [The model of the set of information spaces in the problem of insider detection]. *Informatika i ee Primeneniya — Inform. Appl.* 11(4):65–69. doi: 10.14357/19922264170408. EDN: ZXWUOP.
15. IBM Watson. Available at: <https://www.ibm.com/watson/services/natural-language-understanding> (accessed March 14, 2025).
16. GigaChat. Available at: <https://developers.sber.ru/portal/products/gigachat> (accessed March 14, 2025).
17. YandexGPT. Available at: <https://yandex.cloud/en/services/yandexgpt> (accessed March 14, 2025).

Received January 9, 2025

Accepted February 15, 2025

Contributors

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

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

Zabezhailo Michael I. (b. 1956) — Doctor of Science in physics and mathematics, principal scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; m.zabezhailo@yandex.ru

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

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

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

ИССЛЕДОВАНИЕ ВЛИЯНИЯ МЕХАНИЗМА ЛАПЛАСА НА СТАТИСТИЧЕСКИЕ ХАРАКТЕРИСТИКИ ИСХОДНЫХ ДАННЫХ*

В. Н. Гридин¹, Б. Р. Салем², Д. С. Смирнов³, В. И. Солодовников⁴

Аннотация: Использование дифференциальной приватности (ДП) предполагает добавление контролируемого шума в исходные данные или результаты вычислений, что позволяет обеспечить надежное и доказуемое сохранение конфиденциальности, но способно заметно сказаться на статистических свойствах наборов данных, потенциально влияя на их последующий анализ. Такая особенность вынуждает искать баланс между приватностью и полезностью. Представлены результаты исследования влияния механизма Лапласа, обеспечивающего ДП, на различные распределения данных и их статистические свойства. В рамках трех экспериментов изучены особенности изменения ряда распределений при применении механизма Лапласа с различными значениями бюджета приватности (эпсилон), проведено сравнение классического механизма Лапласа с его расширениями в контексте их эффективности и рассмотрено влияние на различные уровни корреляции признаков в исходных данных. Результаты подчеркивают компромиссы между приватностью и полезностью данных, а также предоставляют рекомендации по выбору подходящих ДП-механизмов для различных сценариев.

Ключевые слова: дифференциальная приватность; механизм Лапласа; информационная безопасность; математическая статистика; корреляция; конфиденциальность; сокрытие данных

DOI: 10.14357/08696527250104

EDN: NOXOPY

1 Введение

В эпоху больших данных сбор и анализ массивов информации стали повсеместными. Несмотря на аналитическую ценность, многие наборы данных содержат требующую защиты конфиденциальную информацию об отдельных лицах. Дифференциальная приватность [1, 2] стала надежной математической основой, позволяющей вводить контролируемый шум в исходные данные или

* Исследование выполнено в рамках государственного задания: FFNR-2024-0003.

¹ Центр информационных технологий в проектировании Российской академии наук, info@ditc.ras.ru

² Центр информационных технологий в проектировании Российской академии наук, isub97@gmail.com

³ Центр информационных технологий в проектировании Российской академии наук, dvonrims@yandex.ru

⁴ Центр информационных технологий в проектировании Российской академии наук, v_solodovnikov@hotmail.com

результаты вычислений и тем самым гарантировать, что наличие либо отсутствие отдельной записи не приведет к существенному изменению выводов. Несмотря на защиту приватности, введение шума может заметно сказываться на статистических свойствах наборов данных, потенциально влияя на их последующий анализ, что вынуждает исследователей искать баланс между приватностью и полезностью. Этот баланс нередко описывается понятием «бюджет приватности» ϵ .

Представленные в работе результаты дополняют существующие исследования путем последовательного сопоставления трех механизмов Лапласа (классического, усеченного и ограниченного) на шести ключевых типах распределений, включая равномерное, нормальное, Стюдента, хи-квадрат, экспоненциальное и Лапласа. Единообразие экспериментов отражается в одинаковом размере выборки, единой метрике и единых критериях. Такой подход дает систематическое представление об искажениях, возникающих при добавлении шума, в том числе о смещениях среднего, росте или падении дисперсии и изменении формы распределения, а также влияние на корреляционные свойства атрибутов выборки.

Таким образом, основной вклад данной работы — систематическая экспериментальная оценка механизмов Лапласа на различных синтетических данных, включающая:

- анализ влияния бюджета приватности (ϵ) на сохранение статистических свойств для различных распределений;
- сравнение эффективности классического механизма Лапласа с его расширениями (усеченным и ограниченным механизмами Лапласа) в отношении полезности данных;
- исследование влияния механизма Лапласа на корреляцию между переменными.

2 Дифференциальная приватность и текущее состояние исследований

Дифференциальная приватность была предложена и рассматривается как перспективный метод сохранения конфиденциальности и ее количественной оценки. Она может обеспечить надежное и доказуемое сохранение конфиденциальности, что отражается в трех аспектах:

- (1) обеспечении строгого и настраиваемого уровня сохранения конфиденциальности, поскольку добавление или удаление любой записи не влияет на результаты;
- (2) защите от модели атаки, поскольку она не зависит от знания того, какими базовыми знаниями обладает злоумышленник;
- (3) обладании полными и доказуемыми математическими теоретическими основами.

По сравнению с такими методами обезличивания данных, как k -анонимность, l -разнообразие и t -близость, которые требуют предположений о конкретных

атаках и базовых знаний, ДП становится все более популярной в теоретической и практической сферах.

Стандартное определение ДП устанавливает мультипликативную верхнюю границу изменения плотности распределения в наихудшем случае. Неформально такой подход к обезличиванию ограничивает сдвиг в распределении выходных данных рандомизированного алгоритма, который может быть вызван небольшим изменением входных значений.

Таким образом, ДП служит условием механизма предоставления доверенной стороной информации о наборе данных, а не самого набора данных. Интуитивно соображение заключается в том, что для любых двух схожих наборов данных ДП-алгоритм будет вести себя практически одинаково. Определение дает надежную гарантию того, что присутствие или отсутствие приватных данных конкретного человека не окажет существенного влияния на конечный результат алгоритма.

Дифференциальная приватность была формализована Дворк [3] и с тех пор стала краеугольным камнем в анализе приватных данных. Многочисленные исследования изучали разработку ДП-механизмов и их применение. Однако меньше внимания уделялось пониманию того, как эти механизмы влияют на базовые статистические свойства данных.

В одном из исследований [4] был проведен анализ статистических последствий ДП с акцентом на то, что добавление шума может значительно повлиять на статистические оценки. Нередко в статьях (см., например, [5]) изучают компромиссы между приватностью и полезностью, подчеркивая важность выбора механизма. Кроме того, в [6] было проанализировано влияние ДП на модели машинного обучения и отмечено ухудшение производительности моделей при усилении ограничений приватности.

Механизм Лапласа считается одним из наиболее широко применяемых инструментов ДП благодаря своей концептуальной простоте и ясным теоретическим гарантиям [7, 8]. На его основе были предложены различные расширения, в том числе усеченный [9], ограниченный [10] и гибридные варианты, позволяющие учитывать особенности распределения данных. Существуют и иные механизмы, например механизм на основе гамма-распределений [11] или комбинированный [12], но их использование предполагает учет дополнительных параметров (например, δ) и иногда оказывается сложнее в реализации [13, 14]. Ранние методы фокусировались на избирательном добавлении шума (лапласовском [2], гауссовском [15] и др.), но за последние несколько лет появились работы, прицельно анализирующие улучшенные и гибридные механизмы. Например, в [12] приводится подход к объединению шума Лапласа и Гаусса, а в [16] рассматриваются механизмы, адаптирующиеся к индивидуальным особенностям данных. При этом подобные исследования [17, 18] зачастую сосредоточены либо на теоретических оценках ошибок и доказательствах (ϵ, δ) -приватности, либо на отдельных моделях, не дающих цельного сравнительного анализа сразу нескольких механизмов на различных распределениях. Многие актуальные исследования, посвященные гибридным методам, подтверждают, что на практике выбор между

классическим и более сложными механизмами по-прежнему сводится к компромиссу между уровнем приватности и приемлемой искаженностью данных [19–21]. По этой причине данная работа сосредоточена именно на семействе лапласовских механизмов, сохраняющих востребованность в прикладных сценариях [22, 23].

В работе [24] были исследованы статистические свойства ограниченного и усеченного механизмов Лапласа, но без анализа ДП ограниченного механизма Лапласа. В последующей работе [15] показано, что для удовлетворения ДП в ограниченном механизме Лапласа требуется удвоение дисперсии шума ($\epsilon = 2\epsilon'$). В исследовании [25], посвященном регрессионному анализу при ДП, авторы добавляли шум (используя механизм Лапласа) к коэффициентам целевой функции, чтобы обеспечить ДП. Однако это приводило к неограниченной целевой функции. Их первым подходом было многократное использование ДП-механизма до получения решения задачи оптимизации. Этот подход также удваивал дисперсию шума ($\epsilon = 2\epsilon'$). Авторы также предложили альтернативный подход к сохранению бюджета приватности на уровне ϵ . Еще одно исследование наивного байесовского классификатора [26] использовало добавление лапласовского шума к параметрам модели для обеспечения ДП. Для числовых данных наивный байесовский классификатор рассчитывал среднее и стандартное отклонение признака для классификации неизвестных данных. Авторы предложили повторную выборку из распределения Лапласа, чтобы гарантировать, что ДП стандартные отклонения будут положительными, не изменяя дисперсию.

Ранее также исследовалась проблема согласованности в ДП. Примеры включают расчет частот и гистограмм. В работе [27] авторы стремились продемонстрировать частоты, состоящие из неотрицательных целых чисел, с согласованными суммами через использование преобразований Фурье и линейного программирования. В работе [28] авторы использовали ограниченный вывод для обеспечения согласованности частот в гистограммах через постобработку.

В свою очередь, настоящее исследование служит практическим дополнением к классическим теоретическим разработкам по ДП и достижениям современных работ, изучающих применение шума в реальных задачах [22, 29]. Результаты проведенного исследования основываются на этих работах, предоставляя всесторонний анализ влияния механизма Лапласа и его расширений на статистические характеристики выбранных распределений, и включают набор тестов, позволяющих оценить степень искажения исходных данных в зависимости от выбранного уровня приватности. Подобный подход позволит создать фундамент для дальнейших исследований различных механизмов ДП в контексте единой платформы валидации, предлагая практические сведения для специалистов по данным.

3 Механизм Лапласа

Механизм Лапласа [2] представляет собой прототип ДП-алгоритма, позволяющий выдавать приблизительный (зашумленный) ответ на произвольный запрос со значениями в $\mathbb{R}^n$. Этот механизм добавляет шум Лапласа (т. е.

шум из распределения Лапласа, который может быть выражен функцией плотности вероятности $\text{noise}(x) \propto \exp(-|x|/\lambda)$, имеющей нулевое среднее значение и стандартное отклонение $\sqrt{2}\lambda$. Тогда для любой функции $f: \mathbb{D} \rightarrow \mathbb{R}^d$ механизм Лапласа применяет шум к результатам f и определяется как

$$A(f(x), \varepsilon) = f(x) + (Z_1, \dots, Z_d),$$

где Z_i — независимые и одинаково распределенные случайные величины, полученные из распределения Лапласа с масштабом $\Delta f/\lambda$, а Δf — ℓ_1 -чувствительность функции f . Можно рассмотреть $\Delta f/\lambda$ как фактор конфиденциальности ε .

Тогда механизм Лапласа можно переписать в следующем виде:

$$\mathbf{L}_\varepsilon f(x) \triangleq f(x) + \text{Lap}\left(0, \frac{\Delta_1 f}{\lambda}\right),$$

где $\text{Lap}(0, \Delta_1 f/\lambda)$ — распределение Лапласа, а $\Delta_1 f$ соответствует ℓ_1 -чувствительности запроса f , равной

$$\Delta_1 f \triangleq \max_{D_1, D_2} \|f(D_1) - f(D_2)\|_1$$

по всем соседним входам D_1 и D_2 .

Рассмотренное в текущей статье семейство лапласовских механизмов представлено следующим набором:

- классический механизм Лапласа [1] — добавляет шум из распределения Лапласа к каждому элементу данных без каких-либо ограничений;
- усеченный механизм Лапласа [9] — добавляет шум из распределения Лапласа, но шум усечен до определенных границ, чтобы предотвратить появление экстремальных значений;
- ограниченный механизм Лапласа [10] — добавляет шум и ограничивает итоговые зашумленные данные определенным диапазоном, соответствующим диапазону исходных данных.

4 Исследование влияния механизма Лапласа на статистические свойства распределений

В данном эксперименте было изучено, как механизм Лапласа влияет на статистические свойства различных распределений при различных значениях параметра приватности ε . Целью исследования ставится возможность определения компромисса между уровнем приватности и сохранением статистической полезности данных.

В табл. 1 представлены распределения и соответствующие параметры, которые использовались для генерации исходных выборок в экспериментах. Выбор

Таблица 1 Таблица распределений, используемых в эксперименте

Распределение	Параметры	Назначение/особенности
Равномерное	Диапазон значений: $[0, 1]$	Ограниченный интервал $[0, 1]$, равномерная плотность, часто используемая базовая модель генерации данных
Нормальное	Коэффициент сдвига и масштаба ($\mu = 0, \sigma = 1$)	Классическое «центральное» распределение (Gaussian), симметричное
Стьюдента	Степени свободы: $df = 10$	Распределение с «тяжелыми» хвостами, часто встречающееся в статистике (t-распределение)
Хи-квадрат	Степени свободы: $df = 2$	Одностороннее (неотрицательное), значимый пример несбалансированного, скошенного распределения
Экспоненциальное	Обратный коэффициент масштаба: $\lambda = 1$	Модель времени между событиями, неотрицательная, с длинным правым хвостом
Лапласа	Коэффициент сдвига и коэффициент масштаба ($\mu = 0, b = 1$)	Распределение Лапласа (двойной экспоненциальной формы), более «острый пик», тяжелые хвосты

обусловлен целью отразить разнообразие и особенности форм распределений (симметричное/скошенное, тяжелые хвосты, ограниченный интервал и т. п.), часто встречающихся в реальных данных.

Для каждого распределения была сгенерирована выборка размером 1000 элементов с использованием соответствующих параметров. При использовании больших наборов данных ($n \gg 1000$) качественные выводы (в каком диапазоне искажения становятся критичными, как меняются гистограммы) будут аналогичными, но количественно колебания (дисперсии оценок) станут меньше. Механизм Лапласа был применен к каждой выборке с различными значениями ε : 0, 1; 0,5; 1; 5; 10. Чувствительность функции (Δf) была рассчитана как размах данных. Для исходных и зашумленных данных были рассчитаны следующие статистические показатели: среднее значение; дисперсия; стандартная ошибка среднего; доверительные интервалы для среднего (95%); коэффициенты асимметрии и эксцесса; различия в средних и стандартных отклонениях между исходными и зашумленными данными. Каждое из перечисленных распределений генерировалось независимо, что определяло отсутствие между ними дополнительной корреляции.

Для исходных и зашумленных данных были применены статистические тесты как для проверки принадлежности выборок к одной генеральной совокупности,

так и для проверки нормальности распределения зашумленных данных. В первом случае использовался критерий согласия Колмогорова–Смирнова (КС-тест), который не требует предположения о нормальности данных и эффективно указывает на различия в эмпирических функциях распределения. Для проверки нормальности распределения зашумленных данных был выбран критерий Шапиро–Уилка (важно, чтобы шум не приводил к приближению распределения к нормальному при больших ε). Также во многих прикладных задачах (построение доверительных интервалов, применение Т-тестов и т. д.) предполагается нормальность распределений. Дополнительно в связи с выявленным влиянием механизмов Лапласа на дисперсию для сопоставления диапазона значений с равномерным распределением был применен критерий хи-квадрат.

Анализ экспериментальных результатов (таблица доступна по ссылке на интернет-ресурсе¹) показал, что для каждого распределения при фиксированном значении параметра приватности ε оценки зашумленного среднего и дисперсии, а также результаты статистических тестов (КС, Шапиро–Уилка и хи-квадрат) существенно зависят от выбранного механизма Лапласа (рис. 1–5). При малых значениях ε (например, при $\varepsilon = 0,1$) классический механизм приводит к высоким значениям дисперсии: для равномерного распределения зашумленное стандартное отклонение может превышать 211 (при исходном значении 0,085), а среднее сдвигается с 0,496 до 0,701. В то же время усеченный и ограниченный механизмы демонстрируют гораздо более скромные изменения (примерно 0,244 и 0,083 соответственно). Аналогичная тенденция наблюдается и для нормального распределения, где при $\varepsilon = 0,1$ дисперсия классического механизма достигает значений свыше 7000, в то время как альтернативные варианты дают значения порядка 8,2 и 2,6. Эти результаты свидетельствуют о том, что при строгой приватности классический механизм «выбрасывает» значения за пределы допустимого диапазона, особенно в распределениях с тяжелыми хвостами (например, Стьюдента или Лапласа). При увеличении ε до умеренных величин (например, 1 или 5) разница между механизмами существенно уменьшается, так как добавленный шум становится менее значительным, хотя некоторые тесты (например, критерий КС) все еще фиксируют отклонения, указывающие на нарушения формы распределения.

В распределениях с большим разбросом или нижними ограничениями (например, хи-квадрат и экспоненциальное) классический механизм при низких ε порождает высокие значения дисперсии, тогда как усеченный и ограниченный варианты существенно ограничивают выбросы. При высоких значениях ε (от 5 до 10) все три механизма сходятся по своим показателям, хотя для распределений с тяжелыми хвостами статистически значимые различия сохраняются. Дополнительно результаты теста хи-квадрат показали, что ограниченный механизм имеет тенденцию преобразовывать распределение в равномерное при низких значениях ε (например, при $\varepsilon = 0,1$ для всех типов, при $\varepsilon = 0,5$ для Лапласа,

¹<https://github.com/designitras/LaplacelImpact>.

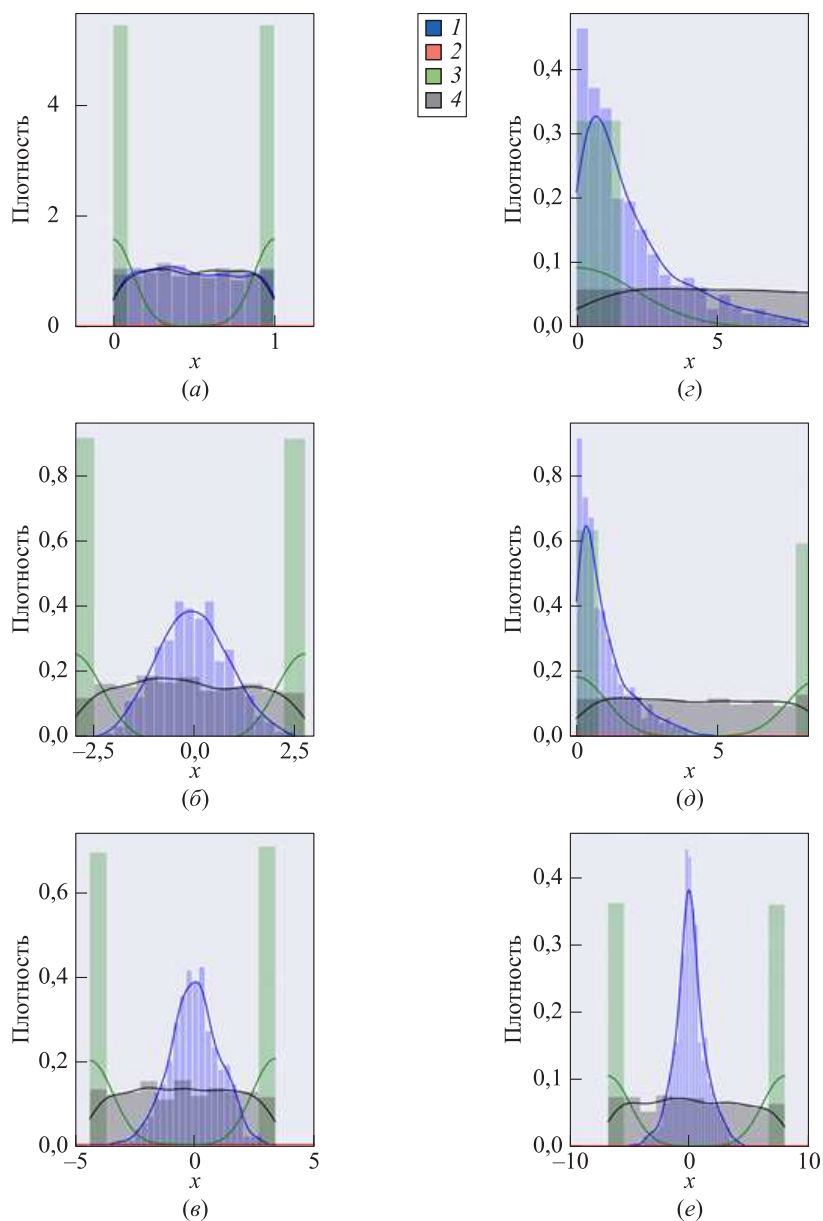


Рис. 1 Визуализация распределений: равномерного (а), нормального (б), Стьюдента (в), хи-квадрат (г), экспоненциального (д) и Лапласа (е) с применением механизмов ДП при $\varepsilon = 0,1$ (1 — исходные данные; 2 — классический Лаплас; 3 — усеченный механизм Лапласа; 4 — ограниченный механизм Лапласа)

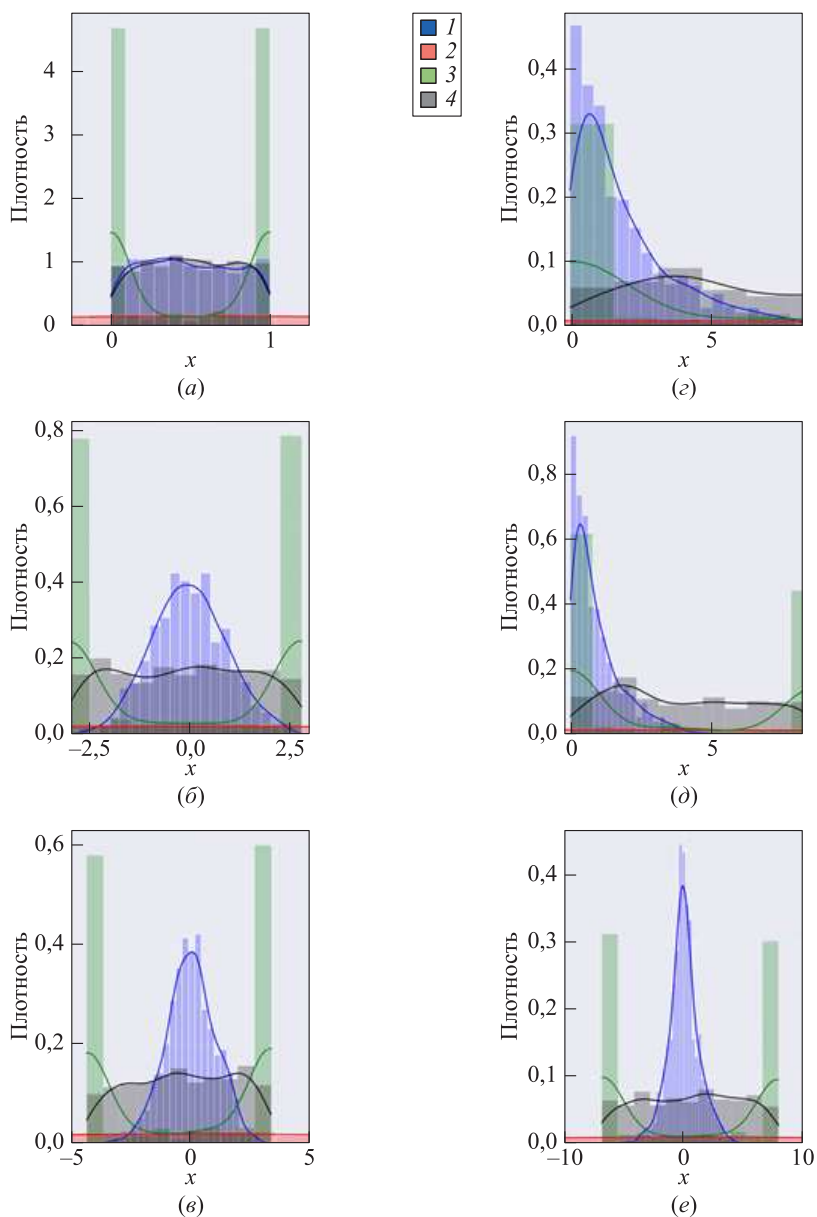


Рис. 2 Визуализация распределений: равномерного (а), нормального (б), Стюдента (в), хи-квадрат (г), экспоненциального (д) и Лапласа (е) с применением механизмов ДП при $\epsilon = 0,5$ (1 — исходные данные; 2 — классический Лаплас; 3 — усеченный механизм Лапласа; 4 — ограниченный механизм Лапласа)

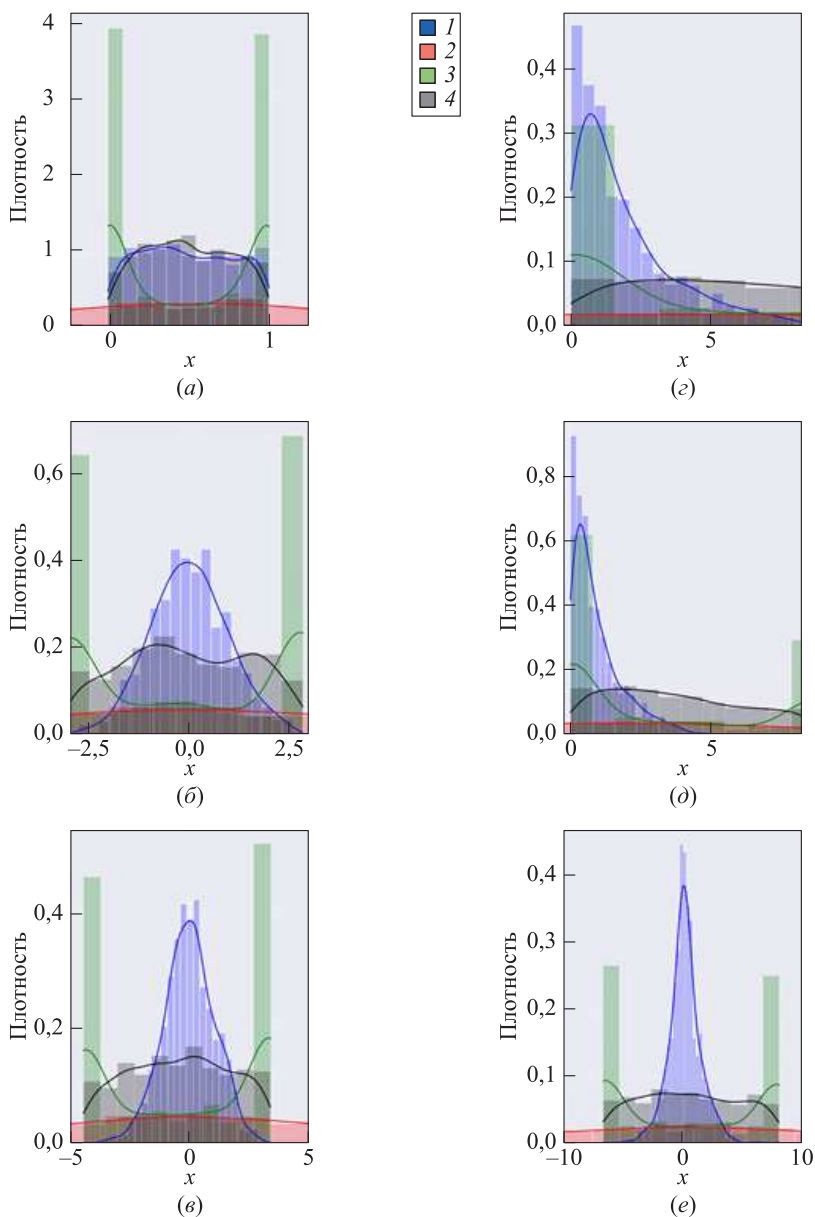


Рис. 3 Визуализация распределений: равномерного (а), нормального (б), Стюдента (в), хи-квадрат (г), экспоненциального (д) и Лапласа (е) с применением механизмов ДП при $\varepsilon = 1$ (1 — исходные данные; 2 — классический Лаплас; 3 — усеченный механизм Лапласа; 4 — ограниченный механизм Лапласа)

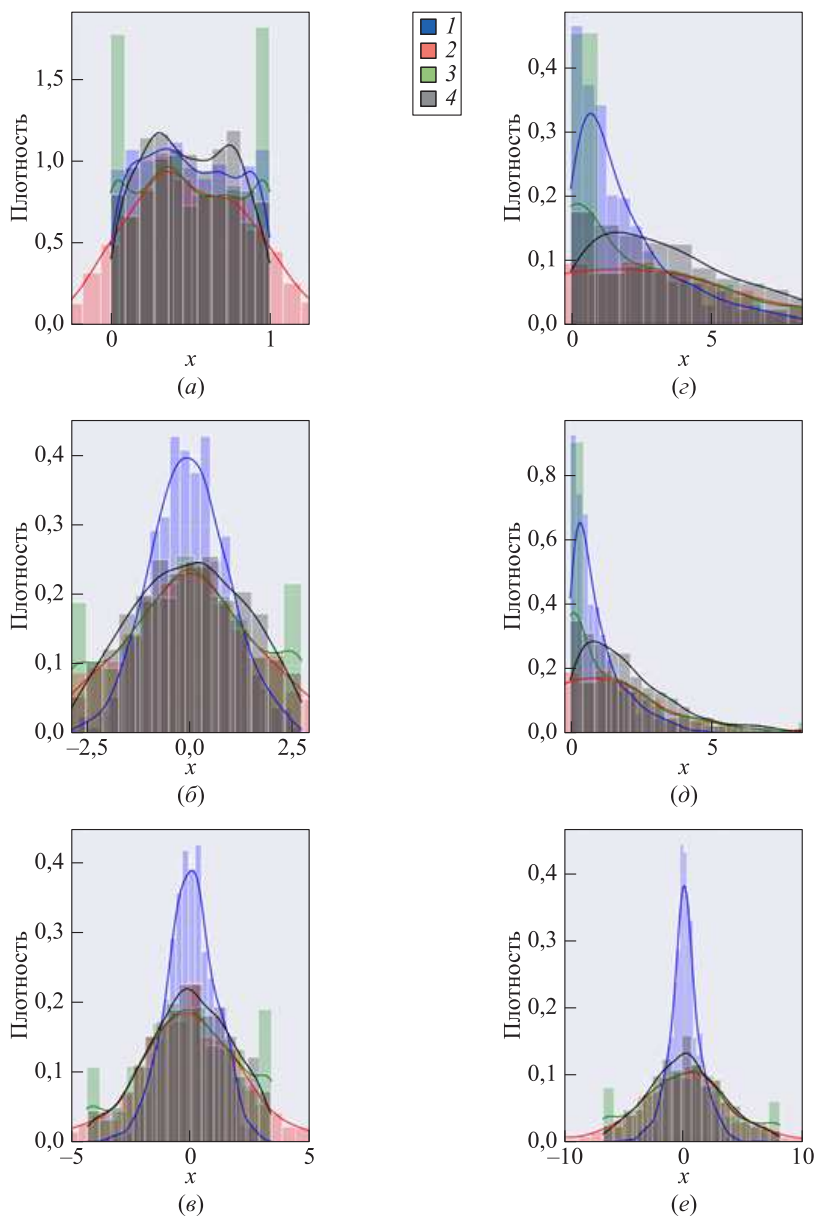


Рис. 4 Визуализация распределений: равномерного (а), нормального (б), Стюдента (в), хи-квадрат (г), экспоненциального (д) и Лапласа (е) с применением механизмов ДП при $\epsilon = 5$ (1 — исходные данные; 2 — классический Лаплас; 3 — усеченный механизм Лапласа; 4 — ограниченный механизм Лапласа)

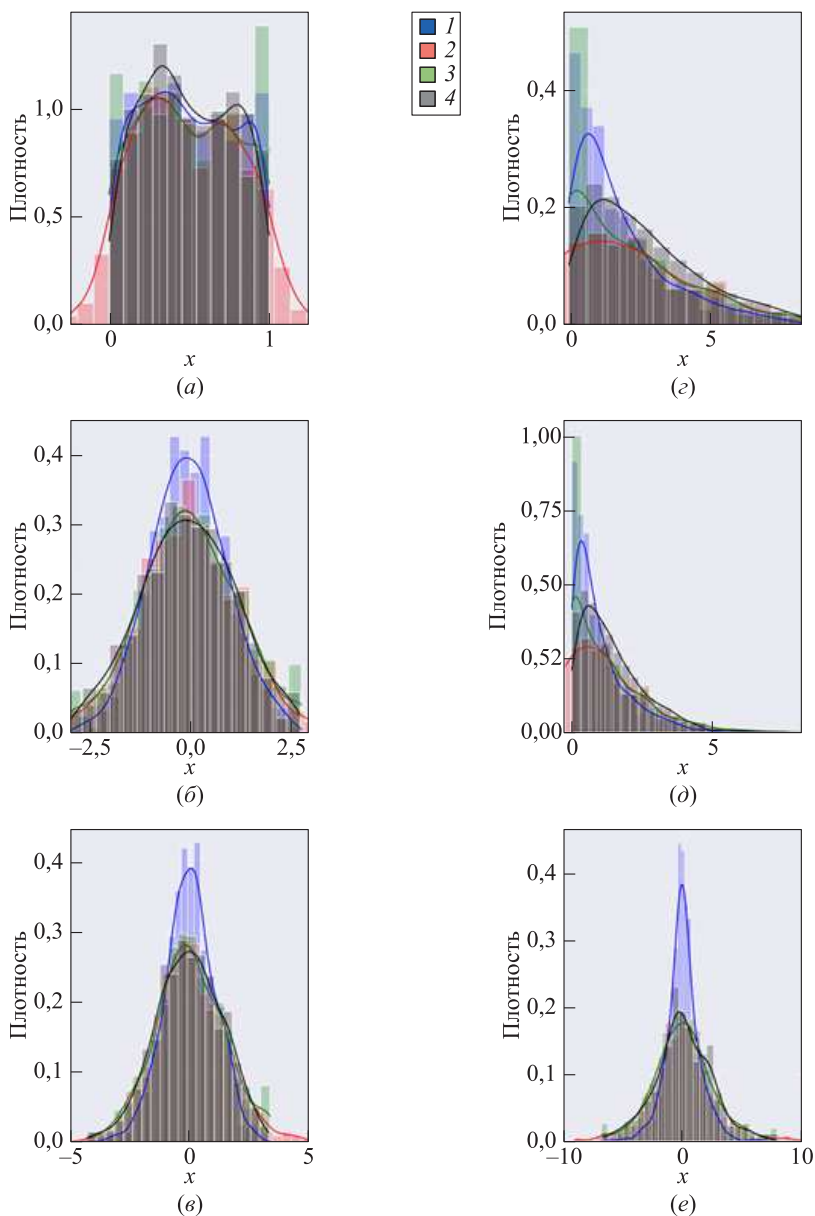


Рис. 5 Визуализация распределений: равномерного (а), нормального (б), Стюдента (в), хи-квадрат (г), экспоненциального (д) и Лапласа (е) с применением механизмов ДП при $\varepsilon = 10$ (1 — исходные данные; 2 — классический Лаплас; 3 — усеченный механизм Лапласа; 4 — ограниченный механизм Лапласа)

экспоненциального и хи-квадрат, а также для нормального распределения при $\varepsilon = 0,1, 0,5, 1,0$ и $5,0$).

В совокупности эти наблюдения подтверждают, что классический механизм при строгой приватности ($\varepsilon < 1$) приводит к искажениям, в то время как усеченный и ограниченный механизмы обеспечивают более стабильное сохранение статистической точности, хотя ограниченный механизм в некоторых случаях может вызывать значительное смещение среднего, если большая часть данных упирается в границы диапазона.

На основе полученных результатов можно сделать следующие выводы. При малых значениях ε тесты КС и Шапиро–Уилка фиксируют значительные отклонения, тогда как при больших ε некоторые механизмы обеспечивают $p\text{-value} > 0,05$, что свидетельствует о приемлемом сходстве с исходными данными. Таким образом, выбор механизма и параметра ε определяется важностью физических ограничений, характеристиками хвостов распределения и требуемым уровнем приватности. При строгой приватности ограниченные и усеченные механизмы, как правило, приводят к меньшим искажениям, хотя усеченный может смещать среднее из-за концентрации значений на границах. При больших ε различия между механизмами практически исчезают, оставаясь лишь незначительными по форме.

5 Исследование влияния различных механизмов Лапласа на среднеквадратичную ошибку при изменении параметра ε

На основе результатов предыдущих разделов было решено дополнительно оценить степень искажения данных с помощью метрик качества. Для этого рассмотрено влияние параметра приватности ε на среднеквадратичную ошибку (MSE — mean squared error) для трех вариантов механизма Лапласа: классического, усеченного и ограниченного. В данном случае MSE будет иметь вид:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2,$$

где y — истинные значения; $\hat{y}$ — зашумленные значения; n — размер выборки.

Основной целью ставилась оценка качества зашумленных данных при изменении параметра приватности — путем измерения среднеквадратичной ошибки между исходными и зашумленными данными при различных вариациях механизма Лапласа, что может помочь при расчете бюджета приватности.

Аналогично предыдущему эксперименту, размер данных был ограничен 1000 элементами. В разд. 3 использовались значения $\varepsilon \in [0,1; 0,5; 1; 5; 10]$, в данном эксперименте был выбран меньший диапазон от 0,1 до 10 с шагом, обеспечивающим 21 значение ε в рассмотренном диапазоне. Чувствительность функции (Δf) была рассчитана как размах данных. Для повышения воспроизводимости результатов для каждого значения ε осуществлялись 100 повторений

Таблица 2 Таблица значений MSE для механизмов Лапласа

ε	Классический	Ограниченный	Усеченный
0,1	198,754	0,164	0,319
0,5	7,95	0,149	0,262
1,0	1,988	0,131	0,208
1,5	0,883	0,115	0,166
2,0	0,497	0,100	0,135
2,5	0,318	0,087	0,111
3,0	0,221	0,076	0,092
3,5	0,162	0,066	0,077
4,0	0,124	0,058	0,065
4,5	0,098	0,050	0,056
5,0	0,080	0,044	0,048
5,5	0,066	0,039	0,042
6,0	0,055	0,035	0,037
6,5	0,047	0,031	0,033
7,0	0,041	0,028	0,029
7,5	0,035	0,025	0,026
8,0	0,031	0,022	0,023
8,5	0,028	0,020	0,021
9,0	0,025	0,018	0,019
9,5	0,022	0,017	0,017
10,0	0,020	0,015	0,016

эксперимента. Для каждого повторения вычислялось MSE между исходными и зашумленными данными, полученные результаты усреднялись и рассматривались как итоговое значение MSE для рассматриваемого ε . Исходные данные были сгенерированы из равномерного распределения на значениях $[0, 1]$. Помимо этого, для усеченного механизма шум ограничивался в диапазоне $[0, 1]$ аналогично диапазону значений распределения, для ограниченного механизма зашумленные данные ограничивались в диапазоне исходных данных [минимальное значение в сгенерированных данных, максимальное значение в сгенерированных данных].

Итоговые результаты демонстрируют, что при $\varepsilon = 0,1$ классический механизм порождает крайне высокое искажение ($MSE \approx 198,75$) из-за выхода шума за пределы $[0, 1]$, в то время как ограниченный и усеченный механизмы имеют значения MSE на уровне $\approx 0,164$ и $0,319$ соответственно (табл. 2). С увеличением ε все механизмы демонстрируют снижение MSE: при $\varepsilon = 1$ классический механизм достигает значения около 2, а при $\varepsilon = 10$ — примерно 0,02, что свидетельствует о снижении влияния шума. Таким образом, при строгой приватности (низкие ε) классический механизм приводит к чрезмерным искажениям, тогда как ограниченные и усеченные варианты обеспечивают более стабильные результаты, хотя в некоторых случаях усеченный механизм может показывать повышенные MSE

из-за скопления значений на границах. Выбор механизма должен основываться на необходимости сохранения физических ограничений данных и требуемом уровне приватности. При сравнении с результатами предыдущего раздела можно заметить, что усеченный механизм при строгой приватности демонстрирует большие значения MSE, чем усеченный механизм, что может быть вызвано скоплением данных на границах рассматриваемого распределения.

6 Исследование влияния механизма Лапласа на коэффициент корреляции при различных уровнях исходной корреляции

В предыдущих экспериментах были рассмотрены независимые выборки различных распределений без явной корреляции между признаками. В данном эксперименте продемонстрировано, как механизм Лапласа влияет на коэффициент корреляции Пирсона между двумя переменными при различных уровнях исходной корреляции. Целью ставится понимание того, насколько сильно изменится корреляция после применения механизма приватности и как это влияет на статистические выводы. Для наглядности потребовалось сгенерировать выборки с заранее заданным уровнем корреляции (ρ). Такого рода выборку проще всего создать при помощи двумерного нормального распределения с матрицей ковариации. Для оценки изменений были выбраны четыре уровня исходной корреляции между переменными X и Y :

- (1) «сильная» корреляция: $\rho = 0,9$;
- (2) «средняя» корреляция: $\rho = 0,5$;
- (3) «слабая» корреляция: $\rho = 0,2$;
- (4) отсутствие корреляции: $\rho = 0,0$.

Для каждого уровня корреляции были сгенерированы выборки размером $n = 1000$ наблюдений. Было использовано многомерное нормальное распределение с заданной матрицей ковариации

$$\text{Cov} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}.$$

Это обеспечивало требуемый уровень корреляции между X и Y .

Использовался фиксированный бюджет приватности $\varepsilon = 10$, данное число было выбрано как значение, при котором величина MSE была близка к нулю во всех вариациях механизма Лапласа. Предполагаемая чувствительность данных была выбрана как размах данных для каждой переменной. К каждой переменной X и Y отдельно применялся механизм Лапласа. Шум генерировался из распределения Лапласа с масштабом $b = S/\varepsilon$. Итоговые зашумленные данные приобретали значения

$$X' = X + \text{Laplace}(0, b); \quad Y' = Y + \text{Laplace}(0, b).$$

Таблица 3 Статистические характеристики исходных и зашумленных данных

Уровень корреляции	Исходная корреляция	Зашумленная корреляция	Z-тест Фишера	P-value
Сильная	0,889227	0,504566	19,264018	0
Средняя	0,504123	0,270290	6,198957	$5,68 \cdot 10^{-10}$
Слабая	0,171436	0,093695	1,767763	$7,71 \cdot 10^{-2}$
Отсутствующая	-0,038935	-0,002482	-0,814327	$4,15 \cdot 10^{-1}$

Далее осуществлялся расчет коэффициентов корреляции:

- исходного коэффициента корреляции — коэффициента корреляции Пирсона между исходными данными X и Y ;
- зашумленного коэффициента корреляции — коэффициента корреляции Пирсона между зашумленными данными X' и Y' .

Использовалось Z -преобразование Фишера для сравнения двух корреляций

$$z = \frac{z_1 - z_2}{\sqrt{2/(n-3)}}.$$

Здесь $z_1 = \arctan(r_1)$; $z_2 = \arctan(r_2)$, где r_1 и r_2 — исходный и зашумленный коэффициент корреляции соответственно. Для оценки статистической значимости различий рассчитывалось P -значение (табл. 3).

Механизм Лапласа добавляет независимый шум к каждой переменной, что снижает степень линейной зависимости. При высоких исходных корреляциях (например, снижение с 0,889 до 0,504) эффект наиболее выражен, а при слабых корреляциях изменения менее критичны. Если исходная корреляция близка к нулю, шум практически не влияет на взаимосвязь. В итоге применение данного механизма может существенно исказить корреляционные структуры, что негативно скажется на аналитике. Поэтому, если цель — минимизировать искажения именно в корреляционных структурах, стоит рассмотреть другие подходы к приватности или гибридные механизмы, специально оптимизированные под сохранение связей между признаками.

Z -тест Фишера демонстрирует, что изменения в коэффициенте корреляции при сильных и средних уровнях корреляции становятся статистически значимыми. При слабой и отсутствующей корреляции изменения не оказываются статистически значимыми.

Эксперимент показал, что применение механизма Лапласа ДП существенно влияет на коэффициент корреляции между переменными, особенно при высоких исходных уровнях корреляции (рис. 6). Добавление шума снижает степень линейной зависимости, что может приводить к статистически значимым изменениям искажений в данных. При планировании использования механизмов ДП важно

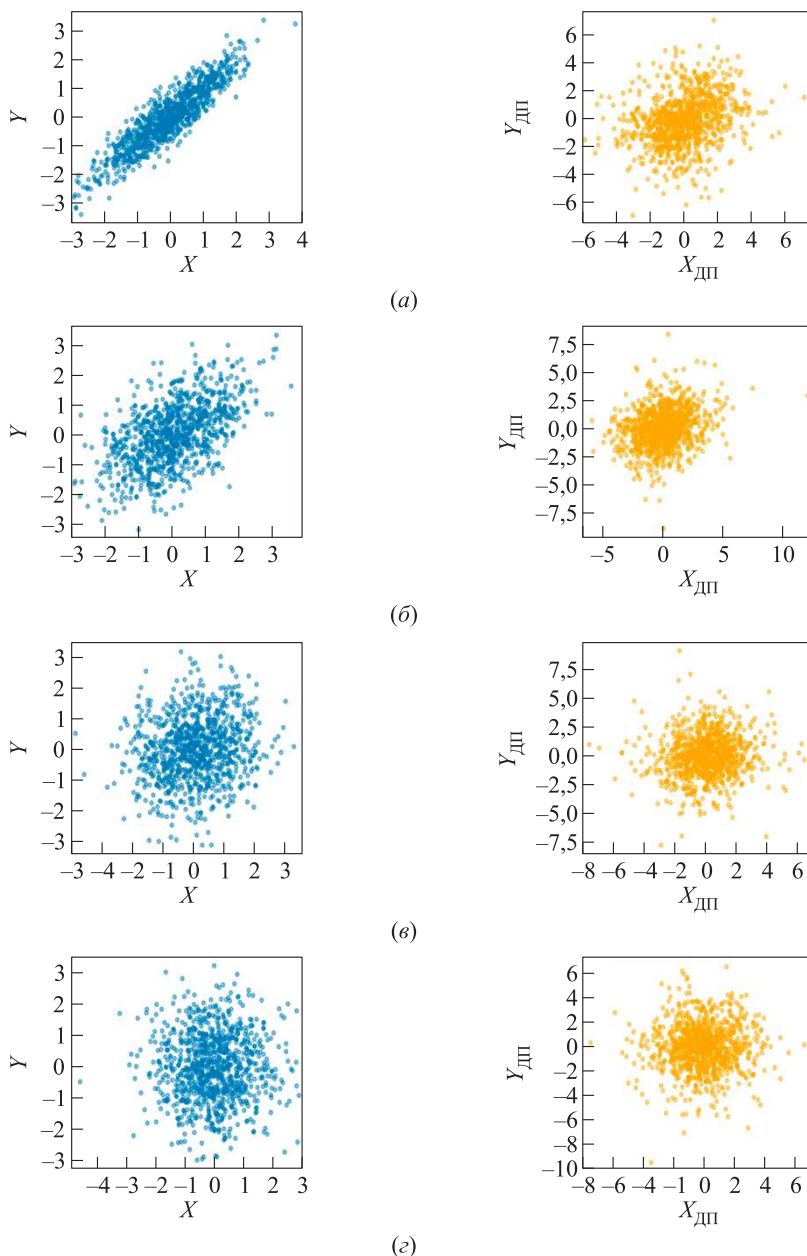


Рис. 6 Диаграммы рассеяния исходных данных (левая колонка) и данных после применения механизма Лапласа (правая колонка) при различных уровнях корреляции: (а) $\rho = 0,9$; (б) $0,5$; (в) $0,2$; (г) $\rho = 0,0$

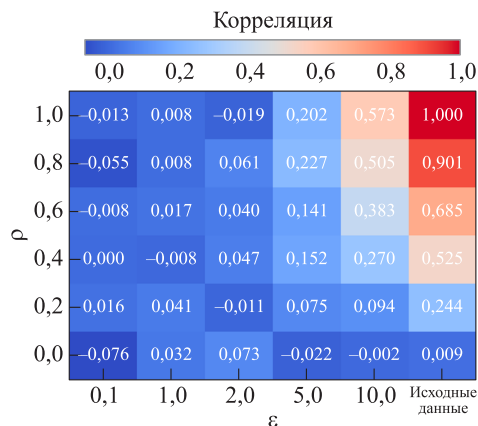


Рис. 7 Матрица корреляций при различных уровнях приватности

реляции остается в пределах значений, демонстрирующих отсутствующую корреляцию независимо от степени корреляции в исходных данных. Данное свойство характеризует невозможность классического механизма Лапласа для сохранения корреляционных свойств при попытке создания частных данных.

7 Заключение

В настоящем исследовании рассмотрено влияние механизма Лапласа на ключевые статистические характеристики данных, включая распределения, среднеквадратичную ошибку и корреляционные зависимости. Результаты показали, что при строгих настройках приватности (низкие значения ϵ) добавляемый шум существенно искажает форму исходных распределений, что приводит к резким изменениям статистических параметров, таких как коэффициенты асимметрии и эксцесса, а также вызывает значительное увеличение MSE. В то же время классический механизм демонстрирует заметное снижение коэффициента корреляции между переменными, что может негативно сказаться на корректности аналитических выводов. Альтернативные варианты — усеченный и ограниченный механизмы — оказывают более умеренное влияние, стабилизируя показатели MSE и лучше сохраняя структуру данных за счет ограничения выбросов. Таким образом, результаты проведенного исследования подчеркивают необходимость тщательного балансирования между уровнем приватности и сохранением статистической полезности данных. Кроме того, полученные выводы указывают на перспективность разработки гибридных методов, способных оптимально сочетать защиту конфиденциальности с минимальными потерями информативности, а также на необходимость дальнейшего изучения влияния механизмов дифференциальной приватности на корреляционные структуры в многомерных

учитывать влияние на корреляционные структуры данных, особенно в задачах, где сохранение зависимостей критично.

Дополнительно для оценки изменения корреляции была составлена матрица сравнений корреляции до и после применения механизма ДП (рис. 7). Были выбраны значения корреляции $\rho = [1,0; 0,9; 0,7; 0,5; 0,2; 0,0]$ и значения фактора конфиденциальности $\epsilon = [10,0; 5,0; 2,0; 1,0; 0,1]$.

Можно заметить, что при значениях $\epsilon < 10$ корреляция зашумленных данных значительно меньше, чем в исходных данных. Помимо этого, в диапазоне $\epsilon = [0,1; 2,0]$ уровень корреляции

выборках. Эти направления представляют значительный интерес для повышения качества аналитических методов в задачах обработки конфиденциальных данных и расширения практического применения дифференциальной приватности.

Литература

1. *Dwork C.* Differential privacy: A survey of results // Theory and applications of models of computation / Eds. M. Agrawal, D. Du, Z. Duan, A. Li. — Lecture notes in computer science ser. — Berlin, Heidelberg: Springer, 2008. Vol. 4978. P. 1–19. doi: 10.1007/978-3-540-79228-4_1.
2. *Dwork C., Roth A.* The algorithmic foundations of differential privacy // Foundations Trends Theoretical Computer Science, 2014. Vol. 9. Iss. 3-4. P. 211–407. doi: 10.1561/04000000042.
3. *Dwork C., McSherry F., Nissim K., Smith A.* Calibrating noise to sensitivity in private data analysis // Theory of cryptography / Eds. S. Halevi, T. Rabin. — Lecture notes in computer science ser. — Berlin, Heidelberg: Springer, 2006. Vol. 3876. P. 265–284. doi: 10.1007/11681878_14.
4. *Wasserman L., Zhou S.* A statistical framework for differential privacy // J. Am. Stat. Assoc., 2010. Vol. 105. No. 489. P. 375–389. doi:10.1198/jasa.2009.tm08651.
5. *Bindschaedler V., Shokri R., Gunter C. A.* Plausible deniability for privacy-preserving data synthesis // Proceedings VLDB Endowment, 2017. Vol. 10. No. 5. P. 481–492. doi: 10.14778/3055540.3055542.
6. *Zhu T., Li G., Zhou W., Yu P.* Differentially private data publishing and analysis: A survey // IEEE T. Knowl. Data En., 2017. Vol. 29. No. 8. P. 1619–1638. doi: 10.1109/TKDE.2017.2697856.
7. *Koufogiannis F., Pappas G. J.* Differential privacy for dynamical sensitive data // 56th Annual Conference on Decision and Control Proceedings. — IEEE, 2017. P. 1118–1125. doi: 10.1109/CDC.2017.8263806.
8. *Muthukrishnan G., Kalyani S.* Differential privacy with higher utility by exploiting coordinate-wise disparity: Laplace mechanism can beat Gaussian in high dimensions. — Cornell University, 2024. arXiv:2302.03511 [cs.CR]. 18 p.
9. *Fu J., Sun Z., Liu H., Chen Z.* Truncated Laplace and Gaussian mechanisms of RDP. — Cornell University, 2023. arXiv:2309.12647 [cs.CR]. 6 p.
10. *Holohan N., Antonatos S., Braghin S., Aonghusa P.* The bounded Laplace mechanism in differential privacy // J. Privacy Confidentialty, 2020. Vol. 10. No. 1. Art. 715. 15 p. doi: 10.29012/jpc.715.
11. *Park Y., Kim M., Yoon J. W.* Differential privacy using gamma distribution // Statistical Signal Processing Workshop Proceedings. — IEEE, 2023. P. 631–635. doi: 10.1109/SSP53291.2023.10207933.
12. *Muthukrishnan G., Kalyani S.* Grafting Laplace and Gaussian distributions: A new noise mechanism for differential privacy // IEEE T. Inf. Foren. Sec., 2023. Vol. 18. P. 5359–5374. doi: 10.1109/TIFS.2023.3306159.
13. *Vadhan S.* The complexity of differential privacy // Tutorials on the foundations of cryptography / Ed. Y. Lindell. — Information security and cryptography ser. — Cham: Springer, 2017. P. 347–450. doi: 10.1007/978-3-319-57048-8_7.

14. Yang M., Chi C.-H., Lam K.-Y., Feng J., Guo T., Ni W. Tabular data synthesis with differential privacy: A survey. — Cornell University, 2024. arXiv:2411.03351 [cs.CR]. 35 p.
15. Liu F. Generalized Gaussian mechanism for differential privacy // IEEE T. Knowl. Data En., 2018. Vol. 31. No. 4. P. 747–756. doi: 10.1109/TKDE.2018.2845388.
16. Huang W., Zhou S., Zhu T., Liao Y., Wu C., Qiu S. Improving Laplace mechanism of differential privacy by personalized sampling // 19th Conference (International) on Trust, Security and Privacy in Computing and Communications Proceedings. — IEEE, 2020. P. 623–630. doi: 10.1109/TrustCom50675.2020.00088.
17. Li X., Luo C., Liu P., Wang L.-E. Information entropy differential privacy: A differential privacy protection data method based on rough set theory // Conference (International) on Dependable, Autonomic and Secure Computing, Pervasive Intelligence and Computing, Cloud and Big Data Computing, and Cyber Science and Technology Congress Proceedings. — IEEE, 2019. P. 918–923. doi: 10.1109/DASC/PiCom/CBDDCom/CyberSciTech.2019.00169.
18. Fioretto F., Van Hentenryck P., Ziani J. Differential privacy overview and fundamental techniques. — Cornell University, 2024. arXiv:2411.04710 [cs.CR]. 23 p.
19. Zhong H., Bu K. Privacy-utility trade-off. — Cornell University, 2022. arXiv: 2204.12057 [cs.IT]. 19 p.
20. Subramanian R. Have the cake and eat it too: Differential privacy enables privacy and precise analytics // J. Big Data, 2023. Vol. 10. No. 1. Art. 117. 14 p. doi: 10.1186/s40537-023-00712-9.
21. Liu Z., Wang W., Liang H., Yuan Y. Enhancing data utility in personalized differential privacy: A fine-grained processing approach // Data security and privacy protection / Eds. X. Chen, X. Huang, M. Yung. — Lecture notes in computer science ser. — Singapore: Springer, 2025. Vol. 15216. P. 47–66. doi: 10.1007/978-981-97-8546-9_3.
22. Cummings R., Desfontaines D., Evans D., et al. Advancing differential privacy: Where we are now and future directions for real-world deployment // Harvard Data Science Review, 2024. Iss. 6.1. 123 p. doi: 10.1162/99608f92.d3197524.
23. Sathish Kumar G., Premalatha K., Uma Maheshwari G., Rajesh Kanna P., Vijaya G., Nivaashini M. Differential privacy scheme using Laplace mechanism and statistical method computation in deep neural network for privacy preservation // Eng. Appl. Artif. Intel., 2024. Vol. 128. Art. 107399. doi: 10.1016/j.engappai.2023.107399.
24. Liu F. Statistical properties of sanitized results from differentially private Laplace mechanism with univariate bounding constraints // Transactions Data Privacy, 2016. Vol. 12. P. 169–195.
25. Zhang J., Zhang Z., Xiao X., et al. Functional mechanism: Regression analysis under differential privacy. — Cornell University, 2012. arXiv:1208.0219 [cs.DB]. 12 p.
26. Vaidya J., Shafiq B., Basu A., Hong Y. Differentially private naive bayes classification // IEEE/WIC/ACM Joint Conference (International) on Web Intelligence and Intelligent Agent Technologies Proceedings. — IEEE, 2013. P. 571–576. doi: 10.1109/WI-IAT.2013.80.
27. Barak B., Chaudhuri K., Dwork C., et al. Privacy, accuracy, and consistency too: A holistic solution to contingency table release // 26th ACM SIGMOD-SIGACT-

- SIGART Symposium on Principles of Database Systems Proceedings. — New York, NY, USA: ACM, 2007. P. 273–282. doi: 10.1145/1265530.1265569.
28. *Hay M., Rastogi V., Miklau G., Suciu D.* Boosting the accuracy of differentially-private histograms through consistency. — Cornell University, 2009. arXiv:0904.0942 [cs.DB]. 15 p.
29. *Das S., Mishra S.* Advances in differential privacy and differentially private machine learning // Information technology security. — Springer Tracts in electrical and electronics engineering ser. — Singapore: Springer, 2024. P. 147–188. doi: 10.1007/978-981-97-0407-1_7.

Поступила в редакцию 27.12.2024

Принята к публикации 15.02.2025

THE LAPLACE MECHANISM IMPACT ON THE STATISTICAL CHARACTERISTICS OF THE ORIGINAL DATA

V. N. Gridin, B. R. Salem, D. S. Smirnov, and V. I. Solodovnikov

Design Information Technologies Center of the Russian Academy of Sciences,
7a Marshal Biryuzov Str., Odintsovo, Moscow Region 143000, Russian Federation

Abstract: Differential privacy (DP) involves adding controlled noise to the input data or computational results which provides robust and provable privacy preservation but can have a significant impact on the statistical datasets properties with potentially influencing effect to their subsequent analysis. This feature forces a trade-off analysis between privacy and utility. The paper presents the research results of the DP Laplace mechanism impact on various data distributions and their statistical properties. The feature changes are studied in the context of three experiments for several key types of distributions while applying the Laplace mechanism with different values of the privacy budget (epsilon). The classical Laplace mechanism is compared with its extensions in the context of their effectiveness and the impact on different levels of feature correlation in the original data is considered. The results highlight the trade-offs between privacy and data utility and also provide recommendations for choosing suitable DP mechanisms for different scenarios.

Keywords: differential privacy; Laplace mechanism; information security; mathematical statistics; correlation; confidentiality; data hiding

DOI: 10.14357/08696527250104

EDN: HOXOPY

Acknowledgments

The study was carried out within the framework of the state assignment: FFNR-2024-0003.

References

1. Dwork, C. 2008. Differential privacy: A survey of results. *Theory and applications of models of computation*. Eds. M. Agrawal, D. Du, Z. Duan, and A. Li. Lecture notes in computer science ser. Berlin, Heidelberg: Springer. 4978:1–19. doi: 10.1007/978-3-540-79228-4_1.
2. Dwork, C., and A. Roth. 2014. The algorithmic foundations of differential privacy. *Foundations Trends Theoretical Computer Science* 9(3-4):211–407. doi: 10.1561/04000000042.
3. Dwork, C., F. McSherry, K. Nissim, and A. Smith. 2006. Calibrating noise to sensitivity in private data analysis. *Theory of cryptography*. Eds. S. Halevi and T. Rabin. Lecture notes in computer science ser. Berlin, Heidelberg: Springer. 3876:265–284. doi: 10.1007/11681878_14.
4. Wasserman, L., and S. Zhou. 2010. A statistical framework for differential privacy. *J. Am. Stat. Assoc.* 105(489):375–389. doi:10.1198/jasa.2009.tm08651.
5. Bindschaedler, V., R. Shokri, and C. A. Gunter. 2017. Plausible deniability for privacy-preserving data synthesis. *Proceedings VLDB Endowment* 10(5):481–492. doi: 10.14778/3055540.3055542.
6. Zhu, T., G. Li, W. Zhou, and P. Yu. 2017. Differentially private data publishing and analysis: A survey. *IEEE T. Knowl. Data En.* 29(8):1619–1638. doi: 10.1109/TKDE.2017.2697856.
7. Koufogiannis, F., and G. J. Pappas. 2017. Differential privacy for dynamical sensitive data. *56th Annual Conference on Decision and Control Proceedings*. IEEE. 1118–1125. doi: 10.1109/CDC.2017.8263806.
8. Muthukrishnan, G., and S. Kalyani. 2024. Differential privacy with higher utility by exploiting coordinate-wise disparity: Laplace mechanism can beat gaussian in high dimensions. Cornell University. 18 p. Available at: <https://arxiv.org/abs/2302.03511> (accessed March 6, 2025).
9. Fu, J., Z. Sun, H. Liu, and Z. Chen. 2023. Truncated Laplace and Gaussian mechanisms of RDP. Cornell University. 6 p. Available at: <https://arxiv.org/abs/2309.12647> (accessed March 6, 2025).
10. Holohan, N., S. Antonatos, S. Braghin, and P. Aonghusa. 2020. The bounded Laplace mechanism in differential privacy. *J. Privacy Confidentiality* 10(1):715. 15 p. doi: 10.29012/jpc.715.
11. Park, Y., M. Kim, and J. W. Yoon. 2023. Differential privacy using gamma distribution. *Statistical Signal Processing Workshop Proceedings*. IEEE. 631–635. doi: 10.1109/SSP53291.2023.10207933.
12. Muthukrishnan, G., and S. Kalyani. 2023. Grafting Laplace and Gaussian distributions: A new noise mechanism for differential privacy. *IEEE T. Inf. Foren. Sec.* 18:5359–5374. doi: 10.1109/TIFS.2023.3306159.
13. Vadhan, S. 2017. The complexity of differential privacy. *Tutorials on the foundations of cryptography*. Ed. Y. Lindell. Information security and cryptography ser. Springer. 347–450. doi: 10.1007/978-3-319-57048-8_7.
14. Yang, M., C.-H. Chi, K.-Y. Lam, J. Feng, T. Guo, and W. Ni. 2024. Tabular data synthesis with differential privacy: A survey. Cornell University. 35 p. Available at: <https://arxiv.org/abs/2411.03351> (accessed March 6, 2025).

15. Liu, F. 2018. Generalized Gaussian mechanism for differential privacy. *IEEE T. Knowl. Data En.* 31(4):747–756. doi: 10.1109/TKDE.2018.2845388.
16. Huang, W., S. Zhou, T. Zhu, Y. Liao, C. Wu, and S. Qiu. 2020. Improving Laplace mechanism of differential privacy by personalized sampling. *19th Conference (International) on Trust, Security and Privacy in Computing and Communications Proceedings*. IEEE. 623–630. doi: 10.1109/TrustCom50675.2020.00088.
17. Li, X., C. Luo, P. Liu, and L.-E. Wang. 2019. Information entropy differential privacy: A differential privacy protection data method based on rough set theory. *Conference (International) on Dependable, Autonomic and Secure Computing, Pervasive Intelligence and Computing, Cloud and Big Data Computing, and Cyber Science and Technology Congress Proceedings*. IEEE. 918–923. doi: 10.1109/DASC/PiCom/CBDCCom/CyberSciTech.2019.00169.
18. Fioretto, F., P. Van Hentenryck, and J. Ziani. 2024. Differential privacy overview and fundamental techniques. Cornell University. 23 p. Available at: <https://arxiv.org/abs/2411.04710> (accessed March 6, 2025).
19. Zhong, H., and K. Bu. 2022. Privacy-utility trade-off. Cornell University. 19 p. Available at: <https://arxiv.org/abs/2204.12057> (accessed March 6, 2025).
20. Subramanian, R. 2023. Have the cake and eat it too: Differential privacy enables privacy and precise analytics. *J. Big Data* 10(1):117. 14 p. doi: 10.1186/s40537-023-00712-9.
21. Liu, Z., W. Wang, H. Liang, and Y. Yuan. 2025. Enhancing data utility in personalized differential privacy: A fine-grained processing approach. *Data security and privacy protection*. Eds. X. Chen, X. Huang, and M. Yung. Lecture notes in computer science ser. Singapore: Springer. 15216:47–66. doi: 10.1007/978-981-97-8546-9_3.
22. Cummings, R., D. Desfontaines, D. Evans, *et al.* 2024. Advancing differential privacy: Where we are now and future directions for real-world deployment. *Harvard Data Science Review* 6(1). 123 p. doi: 10.1162/99608f92.d3197524.
23. Sathish Kumar, G., K. Premalatha, G. Uma Maheshwari, P. Rajesh Kanna, G. Vijaya, and M. Nivaashini. 2024. Differential privacy scheme using Laplace mechanism and statistical method computation in deep neural network for privacy preservation. *Eng. Appl. Artif. Intel.* 128:107399. doi: 10.1016/j.engappai.2023.107399.
24. Liu, F. 2016. Statistical properties of sanitized results from differentially private Laplace mechanism with univariate bounding constraints. *Transactions Data Privacy* 12:169–195.
25. Zhang, J., Z. Zhang, X. Xiao, *et al.* 2012. Functional mechanism: Regression analysis under differential privacy. Cornell University. 12 p. Available at: <https://arxiv.org/abs/1208.0219> (accessed March 6, 2025).
26. Vaidya, J., B. Shafiq, A. Basu, and Y. Hong. 2013. Differentially private naive bayes classification. *IEEE/WIC/ACM Joint Conferences (International) on Web Intelligence and Intelligent Agent Technologies Proceedings*. IEEE. 571–576. doi: 10.1109/WI-IAT.2013.80.
27. Barak, B., K. Chaudhuri, C. Dwork, *et al.* 2007. Privacy, accuracy, and consistency too: A holistic solution to contingency table release. *26th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems Proceedings*. New York, NY: ACM. 273–282. doi: 10.1145/1265530.1265569.
28. Hay, M., V. Rastogi, G. Miklau, and D. Suciu. 2009. Boosting the accuracy of differentially private histograms through consistency. Cornell University. 15 p. Available at: <https://arxiv.org/abs/0904.0942> (accessed March 6, 2025).

29. Das, S., and S. Mishra. 2024. Advances in differential privacy and differentially private machine learning. *Information technology security*. Springer tracts in electrical and electronics engineering ser. Singapore: Springer. 147–188. doi: 10.1007/978-981-97-0407-1_7.

Received December 27, 2024

Accepted February 15, 2025

Contributors

Gridin Vladimir N. (b. 1944) — Doctor of Science in technology, professor, scientific leader, Design Information Technologies Center of the Russian Academy of Sciences, 7a Marshal Biryuzov Str., Odintsovo, Moscow Region 143000, Russian Federation; info@ditc.ras.ru

Salem Basim R. (b. 1997) — scientist, Design Information Technologies Center of the Russian Academy of Sciences, 7a Marshal Biryuzov Str., Odintsovo, Moscow Region 143000, Russian Federation; isub97@gmail.com

Smirnov Dmitry S. (b. 1988) — Candidate of Science (PhD) in economics, head of laboratory, Design Information Technologies Center of the Russian Academy of Sciences, 7a Marshal Biryuzov Str., Odintsovo, Moscow Region 143000, Russian Federation; dvonrims@yandex.ru

Solodovnikov Vladimir I. (b. 1977) — Candidate of Science (PhD) in technology, director, Design Information Technologies Center of the Russian Academy of Sciences, 7a Marshal Biryuzov Str., Odintsovo, Moscow Region 143000, Russian Federation; v_solodovnikov@hotmail.com

ПЕРСПЕКТИВНЫЕ МЕТОДЫ РЕАЛИЗАЦИИ ИНКРЕМЕНТАЛЬНОГО ОБНОВЛЕНИЯ МАТЕРИАЛИЗОВАННЫХ ПРЕДСТАВЛЕНИЙ В СОВРЕМЕННЫХ РЕЛЯЦИОННЫХ СИСТЕМАХ УПРАВЛЕНИЯ БАЗАМИ ДАННЫХ

С. А. Ступников¹, Н. А. Скворцов², Д. О. Брюхов³

Аннотация: Работа посвящена обзору перспективных методов реализации инкрементального обновления материализованных представлений в академических и промышленных реляционных системах управления базами данных (СУБД). Рассмотрены основные направления применений материализованных представлений, общая схема задачи инкрементального обновления, классические методы обновления представлений, базовые промышленные реализации в наиболее распространенных СУБД, реализации в современных СУБД и их особенности. Рассмотрены основные направления академических исследований, их отличия по постановке задачи, подмножеству языка определения представлений, практической или теоретической направленности. Практически ориентированные исследования совершенствуют промышленные СУБД, проводят экспериментальную оценку подхода с различными вариантами масштабирования. Работы теоретической направленности концентрируются на изучении подмножеств языка определения представлений, для которых возможны наилучшие стратегии обновления, и формальном доказательстве сложности алгоритмов. Выделены наиболее перспективные методы и пути реализации в современных СУБД.

Ключевые слова: реляционные системы управления базами данных; материализованные представления; инкрементальное обновление

DOI: 10.14357/08696527250105

EDN: ESRTJA

1 Введение

Материализованные представления — это запросы к базам данных, результаты которых сохраняются и обновляются в СУБД для облегчения доступа к данным в соответствующих базовых структурах данных (таблицах). Пользователи базы данных могут задавать запросы над представлениями или определять представления, в которых уже определенные представления рассматриваются как базовые таблицы.

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

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

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

Поддержка материализованных представлений требует ресурсов. Они не только занимают дополнительное пространство, но и требуют обслуживания: по мере изменения базовых таблиц содержимое материализованного представления устаревает. На практике, когда число измененных записей невелико по сравнению с размером базовых таблиц, лучшей стратегией поддержки материализованных представлений становится инкрементальное обновление (Incremental View Maintenance, IVM) — вычисление и применение к материализованным представлениям только инкрементальных изменений, вызванных изменениями в базовой таблице.

Характерные примеры применения материализованных представлений включают организацию витрин данных, реализацию ETL (extract, transform, and load) процессов в хранилищах данных, реализацию аналитики в реальном времени над потоковыми данными. *Витрины данных* — подмножества хранилищ данных, ориентированные на интересы отдельной структурной части организации или бизнес-направления, могут быть реализованы в виде материализованных OLAP (online analytical processing) запросов с группировкой по множеству измерений и агрегацией по параметрам, подвергаемым анализу [1]. Также материализованные представления могут быть использованы для организации самих *хранилищ данных*. При этом преобразование данных из OLTP-источников осуществляется при помощи нескольких последовательных слоев материализованных представлений. Финальным слоем при этом становится структура хранилища. Материализованные представления могут быть использованы при разработке баз данных в различных предметных областях — от лингвистики [2, 3] до управления землепользованием [4] и нейрофизиологии [5].

Высокую актуальность в современной индустрии приобретают *потоковые данные*, сохраняемые в базах данных как постоянно меняющиеся во времени наборы данных. Характерными примерами служат потоки данных с сенсоров Интернета вещей, данные социальных сетей, данные мониторинга высоконагруженных облачных инфраструктур и т. д. Большое число систем принятия решений опирается на аналитику данных в реальном времени. При этом необходимо постоянное обновление результатов вычислений (агрегация) на новых данных.

Данная статья посвящена обзору современных методов инкрементального обновления материализованных представлений и их реализаций в академических и промышленных системах. Выделены наиболее перспективные методы и пути реализации в современных СУБД.

2 Задача инкрементального обновления представлений и классические методы ее решения

В общей схеме задачи инкрементального обновления материализованного представления на основании *запроса* и *базы данных*, создается *структура дан-*

ных для хранения материализованного представления, соответствующего запросу [6]. Обновление вызывает изменения как в базе данных, так и в структуре данных. При запросе пользователя к данным выходные данные запроса извлекаются по одному кортежу за раз. Временная сложность задачи делится на время предварительной обработки, время обновления и задержку извлечения. *Время предварительной обработки* — это время, необходимое для вычисления структуры данных, которая кодирует выходные данные запроса. *Время обновления* — это время, необходимое для обновления базы данных и структуры данных при вставке или удалении одного кортежа. *Задержка извлечения* — это максимум из трех величин: время между началом извлечения и выводом первого кортежа, время между выводом любых двух последовательных кортежей и время между выводом последнего кортежа и окончанием извлечения. Наилучшей стратегией обновления принято считать такую, при которой достигается константное (относительно запроса и базы данных) время обновления и задержка извлечения.

Классическим считается «чистый» алгебраический метод инкрементального обновления представлений [7]. Изменения в данных базовых таблиц при этом представлены в виде пары так называемых дельта-таблиц, содержащих удаленные строки и вставленные строки соответственно. Обновление представления определяется с помощью набора *равенств распространения изменений*. Каждое равенство определяет для операции языка определения представлений, как «распространять» изменения из одной входной таблицы. По определению представления, выраженному в виде дерева операций, во время компиляции равенства распространения изменений применяются к дереву по принципу «снизу вверх». Начиная с изменений в базовых таблицах, конструируются запросы для вычисления изменений в постепенно увеличивающихся подвыражениях представления и в конечном итоге получаются запросы для вычисления изменений в самом представлении. Сконструированные запросы для вычисления изменений выполняются по расписанию, срабатыванию триггера или команде пользователя. Затем представление обновляется в соответствии с вычисленными изменениями.

Изначально алгебраический подход был разработан для представлений, задаваемых операциями селекции, проекции и соединения с множественной семантикой. В дальнейшем он был расширен на операции декартова произведения, объединения, разности, полусоединения, внешнего соединения, группировки и агрегации с множественной и мультимножественной семантикой.

Алгебраический подход привлекателен по нескольким причинам. Помимо того что он концептуально прост, он модульный и компоуемый: равенства распространения изменений определены отдельно для каждой операции языка определения представлений, и в совокупности они могут обрабатывать сколь угодно сложные запросы. В чистом виде этот подход позволяет выполнять все задачи обновления с использованием только стандартных операций языка запросов, и, следовательно, при этом могут быть эффективно использованы существующие механизмы исполнения и оптимизации запросов. Однако на практике для эффективной реализации обычно требуется отказаться от концептуальной простоты

алгебраического подхода и описывать отдельные аспекты обновления в императивном виде.

Хорошо известные трудности при инкрементальном обновлении представлений заключаются в обработке операций проекции в реляционной алгебре или устранении дубликатов в алгебре мультимножеств, когда часто бывают необходимы дорогостоящие запросы к исходной таблице. Классическая ключевая идея для решения этой проблемы — *подсчет производящих кортежей*. Для каждой строки материализованного представления отслеживается число кортежей базовых таблиц, которые могут произвести эту строку. Это число можно легко поддерживать при изменении исходных таблиц.

Подсчет производящих кортежей полезен не только для поддержания представлений, связанных с проекцией и устранением дубликатов, но и для группирования и агрегации. Следует отметить, что подсчет производящих кортежей требует расширения алгебраического подхода. Основные равенства распространения изменений могут быть неэффективными при их прямом применении, поскольку они не учитывают и не используют дополнительную информацию (например, подсчет производящих кортежей).

Идея подсчета производящих кортежей лежит в основе *алгоритма подсчета* (counting algorithm) [8]. Алгоритм, как и алгебраический подход в целом, использует изменения, внесенные в базовые таблицы, и начальное состояние базовых таблиц для конструирования в качестве выходных данных набора изменений, которые необходимо внести в представление. Этот набор изменений называется *итоговыми дельтами*. В итоговых дельтах вставки представлены положительными значениями подсчета производящих кортежей, а удаления — отрицательными.

Преимущество алгоритма подсчета состоит в учете как вставок, так и удалений за один проход (не требуется отдельно обрабатывать вставки и отдельно — удаления). Алгоритм используется различными коммерческими системами: Oracle [9], DB2, Postgres [10], Redshift [11] — для инкрементального обновления своих материализованных представлений. Алгоритм применим для обновления представлений, определяемых при помощи операций селекции, проекции, соединения, группировки и агрегации [12, 13].

3 Базовые промышленные реализации инкрементального обновления представлений

К базовым промышленным реализациям инкрементального обновления материализованных представлений, основанным на классических подходах, следует отнести реализации в СУБД Oracle [9], PostgreSQL [10] и SQL Server [14].

В СУБД Oracle поддерживаются два способа инкрементального обновления материализованных представлений: быстрое (fast) обновление и синхронное (synchronous) обновление. Быстрое обновление основано на отложенном (deferred) подходе к обновлению представлений, при котором транзакция обновления запус-

кается после завершения транзакции обновления базовой таблицы, что замедляет обновление, но при этом не влечет за собой замедления выполнения операций обновления базовых таблиц. Синхронное обновление основано на немедленном (immediate) обновлении, при котором обновление происходит синхронно с обновлением базовой таблицы.

Быстрое обновление применимо для представлений, созданных как над одной базовой таблицей, так и над несколькими таблицами. В последнем случае поддерживаются следующие виды представлений: материализованные представления, содержащие только операции соединения; материализованные представления, содержащие операции агрегации и соединения; вложенные материализованные представления, созданные над другими представлениями. Быстрое обновление оптимизируется с использованием доступных ограничений первичного и внешнего ключей в столбцах соединения.

Существуют три метода инкрементального обновления: журналы (logs); обновление отслеживания изменений раздела (partition change tracking — PCT); обновление отслеживания изменений логического раздела (logical partition change tracking — LPCT).

Метод с использованием журналов базируется на классических методах инкрементального обновления, включая алгебраический метод и подсчет производящих кортежей [9]. Журнал материализованного представления ведет учет изменений в базовых таблицах, включает информацию о всех добавленных и удаленных строках и о строках, в которых было добавлено, удалено или изменено значение.

Метод обновления PCT используется, когда базовые таблицы физически разделены на разделы (partition). Представление должно иметь такое же партиционирование, как и базовая таблица. Этот метод основан на использовании механизма отслеживания изменений разделов, который позволяет идентифицировать затронутые разделы и/или часть измененных данных в представлении. Метод LPCT похож на PCT, но требует логической, а не физической схемы партиционирования в базовой таблице с использованием диапазонов ключей.

Синхронное обновление — это подход к обновлению таблиц и материализованных представлений в хранилище данных, при котором они обновляются одновременно. В большинстве хранилищ данных таблицы фактов партиционированы по временному измерению, и очень часто инкрементальная загрузка данных состоит в основном из изменений в недавние периоды времени. Синхронное обновление использует эти характеристики для значительного улучшения производительности обновления и пропускной способности.

Синхронное обновление объединяет некоторые элементы методов быстрого обновления на основе журналов и PCT. Запрос, определяющий представление, должен соответствовать схеме «звезда» или «снежинка», таблица фактов и материализованное представление должны быть партиционированы, а ключ раздела материализованного представления должен быть функционально зависимым от ключа раздела таблицы фактов.

В СУБД *PostgreSQL* нет встроенной поддержки инкрементального обновления материализованных представлений, но в 2022 г. появилась первая версия реализации библиотеки *pg_ivm* в виде расширения *PostgreSQL* [10]. В библиотеке был выбран синхронный подход к обновлению представлений, при котором обновление происходит в рамках той же транзакции, что и обновление базовых таблиц. Для этого используются триггеры *AFTER*, добавляемые к базовым таблицам. В настоящее время в библиотеке поддерживаются материализованные представления на основе внутреннего соединения, на основе агрегации с использованием некоторых встроенных операций, с использованием простых подзапросов, включая *EXISTS*-подзапросы и *CTE* (*common table expressions*) подзапросы. Реализация метода инкрементального обновления в библиотеке базируется на классических методах инкрементального обновления, включая алгебраический метод и подсчет производящих кортежей [8].

MySQL Server нет материализованных представлений как таковых, однако реализованы индексированные представления [14]. Добавление уникального кластерного индекса материализует данные, возвращаемые представлением, в этот кластерный индекс, по существу, в отдельную таблицу. Обновление происходит синхронно с изменением данных в таблицах и инкрементально, индексы представлений не перестраиваются полностью при изменении данных, а дополняются или модифицируются. Механизм обновления данных аналогичен механизму обновления данных в индексах обычных таблиц.

4 Реализации инкрементального обновления представлений в современных системах управления базами данных

В большинстве современных СУБД реализуются асинхронные подходы к автоматическому обновлению материализованных представлений. Для этого создаются отложенные транзакции после окончания операции изменения базовых таблиц. В *ClickHouse* [15] используется механизм *MergeTree* быстрой вставки данных по частям в сжатом виде, после которой происходит объединение частей в фоновом режиме. Для инкрементального обновления в сочетании с таким механизмом используются специальные типы агрегации и функции предварительного состояния и слияния, отслеживающие изменения частей. Допускается автоматическое каскадное обновление зависимых материализованных представлений, при котором передаются данные вставленного блока, минуя результат в промежуточных целевых таблицах. В *Redshift* [11, 16] отложенное обновление происходит с учетом текущей загрузки системы, ресурсов, необходимых для обновления, и частоты использования материализованных представлений. *Redshift* также может создавать автоматизированные материализованные представления для оптимизации планирования запросов при повторении группирований или агрегатных функций. В *ClickHouse* и *BigQuery* [17] автоматическое обновление инициируется только при вставке данных в базовые таблицы. Это может быть

связано с накладными расходами обновления существующих данных в сжатом виде. Более того, для соединений в запросе материализованного представления триггеры обновления материализованных представлений срабатывают только при вставке в левую таблицу соединения. Изменение данных в правых таблицах не вызывает срабатывания. Для восстановления целостности данных необходимо пересоздание или полное обновление материализованных представлений. Cassandra, следуя особенностям горизонтальной архитектуры семейства столбцов, всегда автоматически отражает в представлениях изменения базовой таблицы на всех репликах [18].

В ClickHouse, Redshift и BigQuery реализовано периодическое обновление материализованных представлений либо ручное обновление по запросу. При этом ClickHouse при обновлении выполняет запросы на полных данных, а в Redshift и BigQuery возможно как инкрементальное, так и полное обновление.

С асинхронным обновлением, как автоматическим, так и периодическим или ручным, в разных продуктах связаны описанные ситуации, при которых транзакции обновления не имеют доступа к уже обновленным внутри других транзакций данным базовых таблиц. Таким образом поддерживается только конечная целостность данных. В ClickHouse описаны случаи возможной потери данных при ошибках обновления материализованных представлений.

В ClickHouse представлены разные механизмы распределения данных при создании таблиц и материализованных представлений, включая локальные разделы, создание на каждом узле кластера таблиц с одинаковой структурой и механизмы распределенных вызовов к таблицам в кластере. Хранение предварительных состояний и механизмы слияния данных частей задаются в явном виде специальными функциями. В Redshift способ распределения задается при создании таблиц и может указывать репликацию на каждый узел, распределение строк по сегментам узлов в порядке круговой очереди или распределение на основе значений одной из колонок, ключа распределения. Допускается создание материализованных представлений на основе внешних таблиц, сформированных с использованием федеративных запросов. В BigQuery, предназначенном для больших данных, поддерживается распределение данных.

Синхронные подходы к обновлению материализованных представлений в рассмотренных СУБД не поддерживают распределения. Так, в ClickHouse реализован экспериментальный вид локальных материализованных представлений, хранимых в оперативной памяти и обновляемых синхронно и инкрементально, однако в новых версиях и он поддерживаться не будет. В Cassandra изменение реплик материализованных представлений осуществляется синхронно, если они расположены на тех же узлах, что и базовые таблицы. Для обновления реплик между узлами используется формирование дельт с асинхронным обновлением.

Для инкрементального обновления материализованных представлений в ClickHouse, как и для распределения данных, явно используются специальные функции предварительного состояния и слияния, связанные с атрибутами и запросами над ними. Также применяется семейство механизмов, связанных

с таблицами и представлениями и задающих определенные алгоритмы объединения частей данных с предварительным вычислением. Redshift автоматически идентифицирует изменения в базовых таблицах и применяет их к материализованному представлению. Способ идентификации изменений не раскрывается. Полное обновление производится, если инкрементальное обновление невозможно для данного вида запроса. Также некоторые операции могут привести к полному пересчету, даже если запрос поддерживает инкрементальное обновление. В этом случае зависимые материализованные представления помечаются для пересчета при следующем обновлении. BigQuery объединяет данные из кешированного представления с новыми данными, чтобы обеспечить согласованность результатов запроса при использовании материализованного представления. Автоматическое инкрементальное обновление позволяет обновлять только изменившиеся разделы данных. Для каждой из рассмотренных СУБД в документации приведены ограничения используемого подмножества языка SQL для возможности инкрементального обновления.

5 Современные направления академических исследований в области инкрементального обновления

Современные направления академических исследований в области инкрементального обновления материализованных представлений являют собой достаточно пеструю картину. Исследования отличаются по постановке задачи, подмножеству языка определения представлений, практической или теоретической направленности. В разных исследованиях предлагаются разные виды усовершенствований классических подходов.

Практически ориентированными следует считать исследования, совершенствующие промышленные СУБД, нацеленные на реализацию промышленных примеров применения, проводящие экспериментальную оценку подхода с различными вариантами масштабирования. Работы теоретической направленности концентрируются в большей степени на изучении подмножеств языка определения представлений, для которых возможны наилучшие стратегии обновления, и формальном доказательстве сложности алгоритмов. Реализация подходов при этом производится в виде академических прототипов.

Так, для инкрементального обновления материализованных представлений в *бессерверных облачных хранилищах* предложен метод, использующий внешние ключи в соединениях [11]. Мотивация метода основана на том, что подавляющее большинство соединений в аналитических запросах над хранилищами данных использует внешние ключи, при этом классические методы обновления представлений проводят вычисления неоптимальным образом, не учитывая этот факт. Предложен алгоритм, проводящий оптимизированные вычисления обновлений представлений для широкого класса представлений над схемами, включающими формы «звезда» и «снежинка», характерные для хранилищ данных. Реализация

метода проведена в виде расширения в системе Redshift, экспериментальные исследования проведены на кластере с 48 виртуальными процессорами над данными из тестовых наборов TPC-DS и TPC-H объемом до 40 ТБ. Эксперименты показали, что оптимизация повышает эффективность обновления представлений в 2–3 раза.

Для оптимизации запросов в *масштабируемых OLTP-системах, основанных на LSM (Log structured merge) деревьях* (BigTable, Cassandra, Oceanbase и PolarDB), разработаны методы оптимизации обновления представлений, основанные на декомпозиции множественных соединений в последовательность бинарных соединений и асинхронном обновлении представлений [19]. Подход к хранению записей с использованием LSM-деревьев предполагает использование двух видов компонентов: memtable для хранения вставок и обновлений записей в памяти (ориентированный на запись) и sstable для хранения больших данных на диске (ориентированный только на чтение). Архитектура системы включает *основной* (хранящий memtable) и *вспомогательные* серверы (хранящие sstable).

Асинхронное обновление представлений повышает производительность за счет объединения обновлений из нескольких транзакций в одну операцию и уплотнения нескольких обновлений одной и той же строки в одно обновление. Обновление представлений происходит при запросе данных этих представлений. Для представлений создаются компоненты V-memtable и V-sstable. Разработаны два метода обновления, отличающиеся тем, где хранится V-memtable — на основном или на вспомогательных серверах. Один из методов имеет лучшую производительность при низкой нагрузке обновления, другой — при высокой нагрузке обновления.

Реализация методов проведена в виде расширения СУБД Oceanbase, экспериментальные исследования проведены на кластере с 80 физическими ядрами над данными из тестового набора TPC-H. Эксперименты нацелены на исследование влияния аналитических запросов на обработку транзакций, и наоборот — влияния транзакций обновления на производительность аналитических запросов. В обоих сценариях разработанные методы показывают значительное превосходство перед базовым синхронным методом обновления.

Для сценария *поточковой передачи обновлений в базу данных* и обновления представлений для каждого кортежа, когда вспомогательные индексы и материализованные представления могут храниться в *основной памяти* и многократно обновляться без больших затрат, разработаны методы обновления представлений с ациклическими по внешним ключам соединениями [20]. Эта ацикличность означает, что граф соединений таблиц по внешним ключам, построенный по запросу, формирующему представление, ацикличен и имеет ровно одну корневую вершину (в которую не входит ни одно ребро). Идея метода основывается на понятии *живых кортежей*: для листовых таблиц графа соединений все кортежи считаются живыми, для нелистовых отношений кортеж считается живым, если он может соединиться хотя бы с одним живым кортежем из каждого потомка.

Отслеживание живых кортежей корневой таблицы графа позволяет эффективно обновлять представление. Определены структуры индексов, необходимые для обновления, и алгоритмы вставки, обновления и удаления кортежей с учетом поддержки живых кортежей и обновления представления (с константным временем задержки извлечения кортежей). Доказано, что сложность обновления для любой входящей последовательности обновлений линейна от «сложности» последовательности.

Реализация метода проведена в виде расширения системы потоковой обработки данных Flink, экспериментальные исследования проведены на кластере из трех узлов над данными из тестового набора TPC-H. В зависимости от тестового запроса метод показывает преимущество в производительности над конкурентами (системы DYN, DBToaster [21], Trill) от двух до сотен раз.

Для обновления представлений, при определении которых используются *соединения не только по внешним ключам*, для того же сценария использования разработан метод, основанный на замене в плане запроса соединений на композицию полусоединения или пересечения и проекции [22]. Идея метода возникла из обнаружения того факта, что основными источниками суперлинейной сложности обновления представлений служат соединения: если план запроса удастся избавить от соединений, то необходимый объем памяти и время обновления могут быть не более чем линейными. Для представлений, определенных ациклическими по соединениям запросами (*free-connex* запросы), разработаны алгоритмы обновления представлений, сложность которых линейна от «сложности» входящей последовательности обновлений.

Реализация метода проведена в виде расширения системы потоковой обработки данных Flink, экспериментальные исследования проведены на кластере из трех узлов над данными из тестовых наборов SNAP и LDBC-SNB, содержащих большие графы из разных предметных областей. В зависимости от тестового запроса метод показывает преимущество в производительности над конкурентами (системы DBToaster, Trill, Flink SQL) от двух до сотен раз.

Ряд современных академических исследований направлен на поиск *условий наилучшей стратегии обновления* (константное время обновления кортежа, константное время представления результата) для различных специальных видов запросов, используемых для определения представлений (*q-иерархические запросы*, каскадные *q-иерархические запросы*, запросы со свободными образцами доступа — *free access patterns*) [6, 23]. Представления на основе некоторых не *q-иерархических запросов* также могут допускать наилучшую стратегию обновления при условии наличия определенных функциональных зависимостей в базе данных.

6 Заключение

На основании проведенного обзора и анализа существующих подходов к инкрементальному обновлению материализованных представлений в промышлен-

ных и академических системах можно сделать вывод о том, что наиболее перспективный путь их реализации — это расширение существующих (базовых) СУБД на основе академических решений.

Механизмы реализации обновления в современных СУБД, а также многочисленные академические методы (по постановкам задачи обновления, подмножеству языка определения представлений, степени промышленной зрелости) весьма разнообразны. Классические методы инкрементального обновления материализованных представлений, такие как алгебраический метод, метод подсчета производящих кортежей, стали основой для реализации в промышленных СУБД.

Асинхронный подход к обновлению материализованных представлений, реализованный в большинстве рассмотренных СУБД, следует считать основным для распределенного обновления. Асинхронное обновление может быть использовано в разных сценариях в рамках хранилища данных: организации витрин данных, реализации ETL-процессов и аналитики в реальном времени над потоковыми данными. Использование синхронного подхода к обновлению материализованных представлений целесообразно преимущественно для решения задач, выполняемых на одном узле.

При выборе базовой СУБД следует учитывать ее систему хранения и соответствие сценарию использования. Так, LSM-деревья могут обеспечивать высокую скорость записи большого числа отдельных кортежей, и сценарий для их использования — высокая распределенная OLTP-нагрузка на базовые таблицы и аналитические запросы над материализованными представлениями. С другой стороны, структура MergeTree ориентирована на быструю запись больших пакетов данных и чтение данных при аналитических запросах. Ее использование целесообразно при применении представлений для организации витрин данных над хранилищами.

Каскадное обновление вложенных материализованных представлений применимо для организации последовательных слоев материализованных представлений в ETL-процессах.

Наиболее зрелые в промышленном смысле *академические решения* уже реализованы в виде расширений существующих систем. В бессерверных облачных хранилищах с пакетным обновлением эффективна модификация алгебраического подхода и алгоритма подсчета с использованием внешних ключей в соединениях. В масштабируемых OLTP-системах, основанных на LSM-деревьях, когда транзакционная нагрузка сочетается с аналитической нагрузкой на таблицы представлений, эффективна декомпозиция множественных соединений в последовательность бинарных соединений и асинхронное обновление представлений. Для сценария потоковой передачи обновлений в базу данных и обновления представлений для каждого кортежа, когда вспомогательные индексы и материализованные представления могут храниться в основной памяти и многократно обновляться без больших затрат, возможна эффективная реализация обновлений для ациклических по внешним ключам представлений. Существуют отдельные эффективные методы обновления представлений (нацеленные на расширение

поддерживаемого класса запросов), основанные на применении вспомогательных иерархических представлений, переписывании запросов, использовании функциональных зависимостей.

Литература

1. Park C.-S., Kim M.-H., Lee Y.-J. Rewriting OLAP queries using materialized views and dimension hierarchies in data warehouses // 17th Conference (International) on Data Engineering Proceedings. — IEEE, 2001. P. 515–523. doi: 10.1109/ICDE.2001.914865.
2. Loiseau S., Sitchinava D. V., Zalizniak Anna A., Zatsman I. M. Information technologies for creating the database of equivalent verbal forms in the Russian–French multivariant parallel corpus // Информатика и её применения, 2013. Т. 7. Вып. 2. С. 100–109. EDN: QITIKN.
3. Зализняк Анна А., Кружков М. Г. База данных безличных глагольных конструкций русского языка // Информатика и её применения, 2016. Т. 10. Вып. 4. С. 132–141. doi: 10.14357/19922264160414. EDN: XGSIWH.
4. Брюхов Д. О., Ступников С. А. Логическая реляционная модель структур данных для решения задач в предметной области управления землепользованием // Информатика и её применения, 2022. Т. 16. Вып. 4. С. 93–98. doi: 10.14357/19922264220414. EDN: UYAJTD.
5. Брюхов Д. О., Ступников С. А., Ковалёв Д. Ю., Шанин И. А. Архитектура распределенного решения задач анализа данных в области нейрофизиологии // Информатика и её применения, 2021. Т. 15. Вып. 1. С. 78–85. doi: 10.14357/19922264210111. EDN: BFHOXC.
6. Olteanu D. Recent increments in incremental view maintenance // 43rd Symposium on Principles of Database Systems Proceedings. — New York, NY, USA: ACM, 2024. P. 8–17. doi: 10.1145/3635138.3654763.
7. Chirkova R., Yang J. Materialized views // Foundations Trends Databases, 2012. Vol. 4. No. 4. P. 295–405. doi: 10.1561/19000000020.
8. Gupta A., Mumick I. S., Subrahmanian V. S. Maintaining views incrementally // SIGMOD Rec., 1993. Vol. 22. No. 2. P. 157–166. doi: 10.1145/170036.170066.
9. Bello R. G., Dias K., Downing A. R., Feenan J. J., Finnerty J. L., Norcott W. D., Sun H., Witkowski A., Ziauddin M. Materialized views in Oracle // 24th Conference (International) on Very Large Data Bases Proceedings. — San Francisco, CA, USA: Morgan Kaufmann, 1998. P. 659–664.
10. Incremental view maintenance. — PostgreSQL wiki, 2024. https://wiki.postgresql.org/wiki/Incremental_View_Maintenance.
11. Svingos C., Hernich A., Gildhoff H., Papakonstantinou Y., Ioannidis Y. Foreign keys open the door for faster incremental view maintenance // Proc. ACM Manag. Data, 2023. Vol. 1. No. 1. 25 p. doi: 10.1145/3588720.
12. Palpanas T., Sidle R., Cochrane R., Pirahesh H. Incremental maintenance for non-distributive aggregate functions // 28th Conference (International) on Very Large Databases Proceedings. — San Francisco, CA, USA: Morgan Kaufmann, 2002. P. 802–813. doi: 10.1016/B978-155860869-6/50076-7.

13. Gupta H., Mumick I. S. Incremental maintenance of aggregate and outerjoin expressions // Inform. Syst., 2006. Vol. 31. No. 6. P. 435–464. doi: 10.1016/j.is.2004.11.011.
14. Create indexed views. — Microsoft Learn, 2025. <https://learn.microsoft.com/en-us/sql/relational-databases/views/create-indexed-views>.
15. Materialized view. — ClickHouse Docs, 2025. <https://clickhouse.com/docs/en/materialized-view>.
16. Refreshing a materialized view. — Amazon Web Services (AWS), 2025. <https://docs.aws.amazon.com/redshift/latest/dg/materialized-view-refresh.html>.
17. Introduction to materialized views. BigQuery. — Google Cloud, 2025. <https://cloud.google.com/bigquery/docs/materialized-views-intro>.
18. Materialized views. — Apache Cassandra documentation, 2025. <https://cassandra.apache.org/doc/latest/cassandra/developing/cql/mvs.html>.
19. Duan H., Hu H., Qian W., Zhou A. Incremental join view maintenance on distributed log-structured storage // Front. Comput. Sci. Chi., 2021. Vol. 15. 16 p. doi: 10.1007/s11704-020-9310-y.
20. Wang Q., Yi K. Maintaining acyclic foreign-key joins under updates // ACM SIGMOD Conference (International) on Management of Data Proceedings. — New York, NY, USA: ACM, 2020. P. 1225–1239. doi: 10.1145/3318464.338058.
21. Koch C., Ahmad Y., Kennedy O., Nikolic M., Nötzli A., Lupei D., Shaikhha A. DBToaster: Higher-order delta processing for dynamic, frequently fresh views // VLDB J., 2014. Vol. 23. P. 253–278. doi: 10.1007/s00778-013-0348-4.
22. Wang Q., Hu X., Dai B., Yi K. Change propagation without joins // Proceedings VLDB Endowment, 2023. Vol. 16. No. 5. P. 1046–1058. doi: 10.14778/3579075.3579080.
23. Kara A., Nikolic M., Olteanu D., Zhang H. F-IVM: Analytics over relational databases under updates // VLDB J., 2024. Vol. 33. P. 903–929. doi: 10.1007/s00778-023-00817-w.

Поступила в редакцию 14.11.2024

Принята к публикации 15.02.2025

ADVANCED METHODS FOR IMPLEMENTATION OF INCREMENTAL VIEW MAINTENANCE IN MODERN RELATIONAL DATABASE MANAGEMENT SYSTEMS

S. A. Stupnikov, N. A. Skvortsov, and D. O. Briukhov

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

Abstract: The paper reviews advanced methods for implementing incremental view maintenance in academic and industrial relational database management systems (DBMS). The main directions of materialized views application, the general statement of the incremental maintenance problem, classic methods

for view maintenance, basic industrial implementations in the most common DBMS, implementations in modern DBMS, and their distinguishing features are considered. The main directions of academic research, their differences in the problem statement, subsets of the view definition language, and practical or theoretical orientation are considered. Practical-oriented investigations improve industrial DBMS and perform an experimental evaluation of the approach with various scaling options. Theoretical investigations focus on the subsets of the view definition language for which the best update strategies are possible and on the formal proof of the complexity of algorithms. The most promising methods and ways of implementation in modern DBMS are distinguished.

Keywords: relational database management systems; materialized views; incremental maintenance

DOI: 10.14357/08696527250105

EDN: ESRTJA

References

1. Park, C.-S., M.-H. Kim, and Y.-J. Lee. 2001. Rewriting OLAP queries using materialized views and dimension hierarchies in data warehouses. *17th Conference (International) on Data Engineering Proceedings*. IEEE. 515–523. doi: 10.1109/ICDE.2001.914865.
2. Loiseau, S., D. V. Sitchinava, Anna A. Zalizniak, and I. M. Zatsman. 2013. Information technologies for creating the database of equivalent verbal forms in the Russian–French multivariant parallel corpus. *Informatika i ee Primeneniya — Inform. Appl.* 7(2):100–109. EDN: QITIKN.
3. Zalizniak, Anna A., and M. G. Kruzhkov. 2016. Baza dannykh bezlichnykh glagol'nykh konstruksiy russkogo yazyka [Database of Russian impersonal verbal constructions]. *Informatika i ee Primeneniya — Inform. Appl.* 10(4):132–141. doi: 10.14357/19922264160414. EDN: XGSIWH.
4. Briukhov, D. O., and S. A. Stupnikov. 2022. Logicheskaya relyatsionnaya model' struktur dannykh dlya resheniya zadach v predmetnoy oblasti upravleniya zemlepol'zovaniem [Logical relational model of data structures for problem solving in land use management]. *Informatika i ee Primeneniya — Inform. Appl.* 16(4):93–98. doi: 10.14357/19922264220414. EDN: UYAJTD.
5. Briukhov, D. O., S. A. Stupnikov, D. Yu. Kovalev, and I. A. Shanin. 2021. Arkhitektura raspredelennogo resheniya zadach analiza dannykh v oblasti neyrofiziologii [An architecture for distributed data analysis problem solving in neurophysiology]. *Informatika i ee Primeneniya — Inform. Appl.* 15(1):78–85. doi: 10.14357/19922264210111. EDN: BFHOXC.
6. Olteanu, D. 2024. Recent increments in incremental view maintenance. *43rd Symposium on Principles of Database Systems Proceedings*. New York, NY: ACM. 8–17. doi: 10.1145/3635138.3654763.
7. Chirkova, R., and J. Yang. 2012. Materialized views. *Foundations Trends Databases* 4(4):295–405. doi: 10.1561/1900000020.
8. Gupta, A., I. S. Mumick, and V. S. Subrahmanian. 1993. Maintaining views incrementally. *SIGMOD Rec.* 22(2):157–166. doi: 10.1145/170036.170066.

9. Bello, R. G., K. Dias, A. R. Downing, J. J. Feenan, J. L. Finnerty, W. D. Norcott, H. Sun, A. Witkowski, and M. Ziauddin. 1998. Materialized views in Oracle. *24th Conference (International) on Very Large Data Bases Proceedings*. San Francisco, CA: Morgan Kaufmann. 659–664.
10. Incremental view maintenance. 2024. PostgreSQL wiki. Available at: https://wiki.postgresql.org/wiki/Incremental_View_Maintenance (accessed March 5, 2025).
11. Svingos, C., A. Hernich, H. Gildhoff, Y. Papakonstantinou, and Y. Ioannidis. 2023. Foreign keys open the door for faster incremental view maintenance. *Proc. ACM Manag. Data* 1(1):40. 25 p. doi: 10.1145/3588720.
12. Palpanas, T., R. Sidle, R. Cochrane, and H. Pirahesh. 2002. Incremental maintenance for non-distributive aggregate functions. *28th Conference (International) on Very Large Databases Proceedings*. San Francisco, CA: Morgan Kaufmann. 802–813. doi: 10.1016/B978-155860869-6/50076-7.
13. Gupta, H., and I. S. Mumick. 2006. Incremental maintenance of aggregate and outerjoin expressions. *Inform. Syst.* 31(6):435–464. doi: 10.1016/j.is.2004.11.011.
14. Create indexed views. 2025. Microsoft Learn. Available at: <https://learn.microsoft.com/en-us/sql/relational-databases/views/create-indexed-views> (accessed March 5, 2025).
15. Materialized view. 2025. ClickHouse Docs. Available at: <https://clickhouse.com/docs/en/materialized-view> (accessed March 5, 2025).
16. Refreshing a materialized view. 2025. Amazon web services (AWS). Available at: <https://docs.aws.amazon.com/redshift/latest/dg/materialized-view-refresh.html> (accessed March 5, 2025).
17. Introduction to materialized views. BigQuery. 2025. Google cloud. Available at: <https://cloud.google.com/bigquery/docs/materialized-views-intro> (accessed March 5, 2025).
18. Materialized views. 2025. Apache Cassandra documentation. Available at: <https://cassandra.apache.org/doc/latest/cassandra/developing/cql/mvs.html> (accessed March 5, 2025).
19. Duan, H., H. Hu, W. Qian, and A. Zhou. 2021. Incremental join view maintenance on distributed log-structured storage. *Front. Comput. Sci. Chi.* 15:154607. 16 p. doi: 10.1007/s11704-020-9310-y.
20. Wang, Q., and K. Yi. 2020. Maintaining acyclic foreign-key joins under updates. *ACM SIGMOD Conference (International) on Management of Data Proceedings*. New York, NY: ACM. 1225–1239. doi: 10.1145/3318464.338058.
21. Koch, C., Y. Ahmad, O. Kennedy, M. Nikolic, A. Nötzli, D. Lupei, and A. Shaikhha. 2014. DBToaster: Higher-order delta processing for dynamic, frequently fresh views. *VLDB J.* 23:253–278. doi: 10.1007/s00778-013-0348-4.
22. Wang, Q., X. Hu, B. Dai, and K. Yi. 2023. Change propagation without joins. *Proceedings VLDB Endowment* 16(5):1046–1058. doi: 10.14778/3579075.3579080.
23. Kara, A., M. Nikolic, D. Olteanu, and H. Zhang. 2024. F-IVM: Analytics over relational databases under updates. *VLDB J.* 33:903–929. doi: 10.1007/s00778-023-00817-w.

Received November 14, 2024

Accepted February 15, 2025

Contributors

Stupnikov Sergey A. (b. 1978) — Candidate of Science (PhD) in technology, leading scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; sstupnikov@ipiran.ru

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

Briukhov Dmitry O. (b. 1971) — Candidate of Science (PhD) in technology, senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; dbriukhov@ipiran.ru

ИНТЕГРАЦИЯ ЭЛЕКТРОННОГО СЛОВАРЯ С ТЕКСТАМИ ПАРАЛЛЕЛЬНОГО КОРПУСА: НОВЫЙ ТЕОРЕТИЧЕСКИЙ ПОДХОД*

Д. О. Добровольский¹, И. М. Зацман²

Аннотация: Рассматривается проблема интеграции электронного словаря с текстами параллельного корпуса в рамках лексикографической информационной системы (ЛГИС), включающей три компонента: (1) электронный многоязычный словарь; (2) корпус как хранилище параллельных текстов; (3) базу данных (БД) аннотированных переводных соответствий (АПС) и две базы знаний. Предлагаемый подход к интеграции обсуждается на примере немецко-русского словаря и представляет собой синтез ряда концептуальных положений, таких как применение принципа многоуровневой структуризации текстов словарных статей, формирование АПС для многозначных слов и устойчивых словосочетаний и их переводов, установление связей между электронным словарем и хранилищем параллельных текстов по значениям многозначных слов и устойчивых словосочетаний. До настоящего времени подобные ЛГИС создавались только для одноязычных словарей с формированием связей только по леммам. Цель статьи — описать предлагаемый подход к интеграции многоязычного электронного словаря с текстами параллельного корпуса как теоретическую основу создания ЛГИС.

Ключевые слова: лексикографическая информационная система; параллельные тексты; электронный многоязычный словарь; корпус; база данных аннотированных переводных соответствий

DOI: 10.14357/08696527250106

EDN: XWJHVC

1 Введение

В работе [1] дано описание подхода к интеграции, предназначенного для пополнения одноязычного³ электронного словаря новыми статьями, подготовленными на основе текстов корпусов проекта DWDS (Digitales Wörterbuch der

*Исследование выполнено в ФИЦ ИУ РАН за счет гранта Российского научного фонда № 24-18-00155, <https://rscf.ru/project/24-18-00155>.

¹Институт русского языка Российской академии наук; Институт языкознания Российской академии наук; Федеральный исследовательский центр «Информатика и управление» Российской академии наук, dm-dbrv@yandex.ru

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

³Одноязычный словарь, толковый или энциклопедический, содержит словарные статьи на одном языке, которые описывают значения слов и приводят примеры их использования или же — в случае энциклопедических словарей — раскрывают суть обозначаемых этими словами понятий.

deutschen Sprache) Берлинско-Бранденбургской академии наук [2, 3]. Предлагаемый подход дает возможность целенаправленно добавлять новые словарные статьи. Собранный статистика работы пользователей со словарем показала, что каждый шестой запрос дает нулевую выдачу, так как статьи с заданным в запросе заглавным словом отсутствуют в словаре. На основе подхода проекта DWDS разработана человеко-машинная технология формирования новых словарных статей [1].

С помощью этой технологии автоматически порождается «костяк» словарной статьи, где имеется морфосинтаксическая зона, фонетическая транскрипция и примеры, но отсутствует толкование значения(й) заглавного слова. Для лексикографа иметь машинный результат как черновик словарной статьи весьма полезно. Однако при этом отмечено, что хороший «костяк» словарной статьи получилось сформировать только в 20% случаев [1].

Для включения новых корпусных примеров в уже существующие словарные статьи в одной из работ группы DWDS [4] предложен способ полуавтоматического извлечения отдельных предложений, иллюстрирующих функционирование слова в тексте из корпусов проекта DWDS. В этом проекте реализована возможность перехода от леммы словарной статьи к примерам из разных корпусов большого объема, интегрированных в структуру этого проекта, что и позволяет извлекать новые примеры использования исследуемых слов. Большие электронные корпуса текстов предоставляют возможности также и для исследования изменений частоты употребления отдельных значений многозначных слов с течением времени [5].

Три перечисленные задачи (пополнение словаря новыми статьями и примерами использования слов, исследование значений слов и мониторинг эволюции частотностей) решались с использованием разных средств информатики. Основное отличие многоаспектной интеграции, предлагаемой в статье, состоит в том, что в рамках одной и той же ЛГИС решается широкий спектр частных задач на основе нового теоретического подхода. Более широкий спектр функций, реализуемых в ЛГИС, обусловлен следующими ее тремя базовыми возможностями. Во-первых, ЛГИС ориентирована на два и более языков, в то время как рассмотренные выше задачи — только на один язык. Во-вторых, кроме описания слов и их различных значений ЛГИС даст возможность искать в корпусе и описывать в словарных статьях устойчивые конструкции, в том числе многозначные, взятые в своих конкретных значениях. И в-третьих, так как в электронном словаре ЛГИС, представленном в форме словарной БД, описание значения слова или конструкции может со временем изменяться, в ЛГИС фиксируется эволюция описаний их значений во времени.

Цель статьи — описать предлагаемый подход к интеграции многоязычного¹ электронного словаря с текстами параллельного корпуса как основу создания

¹Многоязычный словарь содержит словарные статьи на двух или более языках, которые описывают значения слов, предлагают соответствующие им переводы на один (например, с немецкого на русский в немецко-русском словаре) или более языков и приводят примеры использования слов и их переводов.

ЛГИС, обеспечивающей три перечисленные базовые возможности, и продемонстрировать реализуемость этого подхода на примере немецко-русского электронного словаря.

2 Компоненты лексикографической информационной системы

Предлагаемый новый подход к многоаспектной интеграции реализуется в рамках создания трехкомпонентной ЛГИС, сочетающей в себе: (1) многоязычный электронный словарь (= словарная БД); (2) корпус параллельных текстов; (3) БД, создаваемую на основе надкорпусных БД, концепция которых была предложена М. Г. Кружковым [6–10]. Включение в состав ЛГИС электронного словаря обуславливает необходимость решения трудно формализуемой задачи структуризации словарных статей (как наследуемых информационных ресурсов ЛГИС) на двух или более языках.

Описание постановки и решения задачи структуризации заслуживает отдельной статьи. Для иллюстрации предлагаемого подхода к интеграции отметим здесь только то, что в традиционной лексикографии [11] основное внимание уделяется, как правило, только *первому уровню* структуризации словарной статьи на зоны (леммы, этимологии, фонетической транскрипции, грамматической информации о слове в целом, помет, относящихся к описываемому слову в целом и его отдельным значениям, самих значений этого слова и др.). Однако для включения наследуемых статей в словарную БД ЛГИС требуется предварительно решить задачу многоуровневой структуризации зон на поля словарной БД. Таблица 1 из [12] показывает второй, третий и четвертый уровни структуризации только зоны словарной статьи с описанием значений — другие зоны словарной статьи в этой работе не рассматривались (первый уровень — это традиционное деление всей словарной статьи на зоны).

Для иллюстрации этих трех уровней приведем содержание наследуемой словарной статьи многозначного слова *sollen* (см. рисунок), которую необходимо загрузить в словарную БД ЛГИС после многоуровневой структуризации ее зон. На рисунке показан сокращенный вариант словарной статьи для глагола *sollen*. Значения этого слова пронумерованы; в этом варианте статьи приведены краткие описания первых трех из 13 значений этого слова.

На рисунке серым цветом выделен текст для заполнения в табл. 1 полей строк «Варианты перевода... с комментариями», зеленым цветом — для заполнения поля «Комментарий к слову...», голубым цветом — для заполнения полей строк «Примеры употребления слова...» (эти строки сначала делятся на примеры, а потом каждый пример делится на оригинальный текст и его перевод) и коричнево-зеленым цветом — текст «**man sollte meinen, dass. . .** можно было бы подумать, что. . .» для заполнения полей строки «**Устойчивые конструкции** с использованием соответствующего слова в значении 2».

Кроме того, в этом примере красным цветом выделен текст для заполнения полей «Толкование слова в данном подзначении» и желтым цветом — для

sollen *vtod* (*perf* hat sollen, в неполных предложениях, где пропущен инфинитив полнозначного глагола hat gesollt) **1.** должен, следует; сто́ит, хотелось бы (что-л. делать по чьему-л. указанию, по закону, по правилам, нормам и т. п., а также по собственному желанию); ich soll heute noch in die Stadt fahren я должен сегодня ещё поехать в город; er sollte dreißig Euro bezahlen с него потребовали тридцать евро; **тж. в качестве совета, рекомендации (в форме praet conj)** Sie sollten mit dem Ráuchen áufhören вам следует [сто́ит] бросить ку́рить; **в контекстах, предполагающих просьбу или требование совершить некоторое конкретное действие:** ich sollte ihm die Zeitung mitbringen он проси́л меня́ прине́сти ему́ газе́ту; **в придаточных дополнительных тж. в формах conj:** er sagte, ich sollte lieber heute noch in die Stadt zurückkehren он сказа́л, что мне лу́же (бы́ло) бы верну́ться в го́род ещё сего́дня ... **2.** **в формах pqr conj** как указание на несостоявшееся действие сле́довало бы, сто́ило бы, ну́жно бы́ло бы, должен был бы (по мнению говорящего); ich hätte Ihnen überhaupt nichts sagen ~ зря я вам всё э́то рассказа́ла; **в формах praet conj,** как правило, с частицами eigentlich, doch, указывает на ожидаемое положение вещей, которое не имеет места в соответствующий момент er sollte eigentlich schon geschlafen haben я думал, что он уже спал ... □ **man sollte meinen, dass...** мо́жно бы́ло бы поду́мать, что... **3.** **служит для указания на получение информации от других лиц; при переводе на русский язык предложение начинают словами говорят, полагают и т. п.;** sie soll sehr schön gewesen sein говорят, она́ была́ о́чень краси́ва ...

Пример словарной статьи для глагола sollen (сокращенный вариант)

заполнения полей строки «Примеры употребления слова в данном подзначении».

Важно подчеркнуть, что при первичном формировании словарной БД заполняются далеко не все поля, перечисленные в табл. 1 (например, в приведенном примере для слова в значениях 1 и 3 фразем нет, но они могут появиться потом в результате работы лексикографов с корпусными данными ЛГИС).

В табл. 1 показано, что на втором уровне структуризации зону значений надо сначала разделить на *N составляющих*.

Структура третьего уровня включает пять полей описания каждого значения слова (грамматическая информация о слове в целом (а иногда и в отдельных его значениях, если смена значения влияет на грамматические характеристики),

Таблица 1 Второй, третий и четвертый уровни структуризации словарной статьи [12]

Зона значений	Значение 1	Грамматическая информация о слове в значении 1	
		Стилистические пометы, относящиеся к слову в значении 1	
		Комментарий к слову в значении 1	
		Варианты перевода слова в значении 1 с комментариями	
		Примеры употребления слова в значении 1	
		Подзначение 1	Грамматическая информация о слове в данном подзначении
			Стилистические пометы, относящиеся к слову в данном подзначении
			Комментарий к слову в данном подзначении
			Варианты перевода слова в данном подзначении с комментариями
			Примеры употребления слова в данном подзначении
		Подзначение 2	...
		...	...
		Подзначение M_1	...
	Устойчивые конструкции с использованием слова в значении 1		
	...	...	
	Значение N	...	
		Подзначение 1	...
		Подзначение 2	...
		...	...
		Подзначение M_N	...
	Устойчивые конструкции с использованием слова в значении N		

Примечание. Таблица разработана А. А. Гончаровым.

стилистические пометы, относящиеся к слову в целом, комментарии (если в них есть необходимость уже на уровне описания слова в целом), варианты перевода слова в каждом из конкретных значений, примеры употребления слова в этом значении). Четвертый уровень содержит поля для описания M_i подзначений для каждого из N значения слова ($i = \overline{1, N}$).

На четвертом уровне структуризации разные варианты перевода слова и примеры не распределены по отдельным полям (в табл. 1 они не выделены в отдельные ячейки, т. е. четвертый уровень структуризации не является завершающим).

Необходимость решения задачи многоуровневой структуризации зон словарных статей обусловлена одной из целей интеграции: дать возможность лексикографам добавлять и обновлять словарные статьи на уровне значений языковых единиц и примеров их переводов, а пользователям — обращаться непосредственно из словаря к контекстному использованию слов в корпусе, в том числе по

значениям отдельных слов или устойчивых конструкций. В процессе обновления в словарные статьи лексикографы могут включать и новые значения заглавного слова статьи, и новые варианты перевода слова в его значениях, и новые примеры его употребления в конкретных значениях, извлеченных ими из регулярно пополняемого корпуса текстов и обработанных лексикографами (аутентичные примеры из корпуса, как правило, должны модифицироваться для их соответствия формату словаря). Для добавления новых значений, новых вариантов и примеров перевода необходимо заранее предусмотреть в структуре словарной БД соответствующие поля.

Второй компонент ЛГИС — это регулярно пополняемый корпус как хранилище параллельных текстов, описанию которого посвящены работы [13–15]. Отметим, что в корпусе тексты хранятся в неизменном виде, т. е. для ЛГИС это пополняемый, но не редактируемый информационный ресурс.

Третий компонент ЛГИС включает БД аннотированных переводных соответствий. Он также имеет в своем составе базу индивидуальных знаний, которая заполняется лексикографами персонально, и базу коллективных знаний, которая заполняется ими по результатам согласования индивидуальных описаний новых значений слов или устойчивых конструкций. Другими словами, основой согласования служит индивидуальное лингвистическое знание, полученное лексикографом при анализе параллельных текстов ЛГИС (модель извлечения знания лексикографами ЛГИС, разработанная на основе цифровой спиральной модели [16], приведена на рис. 4 в работе [17]).

Описание нового значения, согласованного в процессе обсуждения, записывается и в базу коллективных знаний, и в электронный словарь ЛГИС. При этом сохраняется связь описания нового значения слова (или устойчивой конструкции) с тем текстовым фрагментом параллельного корпуса, из которого оно было извлечено лексикографами. Если электронный словарь содержит только последнюю версию описания нового значения (в словаре раскрытие значения идет чаще всего на основе анализа сочетаемости слова и способов его перевода), то в третьем компоненте ЛГИС хранятся еще и предыдущие варианты его описания, что дает возможность исследовать динамику лингвистического знания лексикографов во времени [18]. При этом варианты описания значений формируются в базах индивидуальных и коллективных знаний по определенным правилам¹.

3 Интеграция компонентов: концептуальные положения

Предлагаемый подход к интеграции электронного словаря и текстов параллельных корпусов предполагает совместное использование трех компонентов, рассмотренных в предыдущем разделе. В примере, который демонстрирует реализуемость предлагаемого подхода, эти компоненты формируются на основе уже

¹Операции и правила формирования (вывода) вариантов описания значений, разработанные А. А. Гончаровым, описаны в [9, 19].

существующих (наследуемых) лингвистических ресурсов. Электронный словарь создается на основе файлов с текстами немецко-русских статей (см. вышеприведенный пример сокращенного варианта словарной статьи для глагола *sollen*). Корпус ЛГИС формируется на основе параллельных текстов Национального корпуса русского языка. База данных АПС формируется на основе БД немецких модальных глаголов, ранее разработанной в ФИЦ ИУ РАН.

Предлагаемый подход к интеграции представляет собой синтез следующих концептуальных положений:

- (1) применение принципа многоуровневой структуризации текстов наследуемых словарных статей на поля словарной БД ЛГИС;
- (2) пополнение словарной БД новыми статьями, сформированными на основе данных корпуса параллельных текстов;
- (3) пополнение уже существующих в словарной БД структурированных словарных статей описаниями новых значений фокусных многозначных (ФМЗ)¹ слов, фразем и идиом², извлеченных из параллельных текстов;
- (4) лингвистическое аннотирование в базе данных АПС значений ФМЗ слов и фразем/идиом, извлеченных из параллельных текстов (т. е. формирование АПС);
- (5) применение принципа формирования связей между словарной БД и корпусом ЛГИС не только по ФМЗ-словам и фраземам/идиомам, но и по их значениям;
- (6) отражение в базах индивидуальных и коллективных знаний траекторий изменений в описаниях значений ФМЗ-слов и фразем/идиом, извлеченных из параллельных текстов.

Из шести перечисленных положений в статье остановимся на четвертом и пятом, так как именно их реализация отличает предлагаемый подход от существующих отечественных и зарубежных подходов. Важно подчеркнуть, что употребления ФМЗ-слов и фразем/идиом в текстах параллельного корпуса во всем их смысловом многообразии необходимо сопоставить в рамках ЛГИС с релевантными значениями в статьях словарной БД. Центральной оказывается при этом проблема их многозначности. Если слова, обладающие только одним значением, позволяют постулировать связь в ЛГИС между словарной статьей и релевантными корпусными примерами на уровне леммы, то для ФМЗ-слов и фразем/идиом требуется установить связи на уровне значений.

¹Прилагательное «фокусные» обозначает те многозначные языковые единицы ЛГИС, для которых словарь и тексты корпуса связываются не только по лемме, но и по отдельным значениям соответствующего слова или устойчивой конструкции с этой леммой.

²В словарной БД ЛГИС устойчивая конструкция может быть или идиомой, или фраземой. Основное отличие между ними состоит в следующем: в каком именно значении то или иное слово употреблено в составе идиомы, сказать практически невозможно, а в составе фраземы есть опорное слово, которое сохраняет в ней одно из своих словарных значений.

Таблица 2 Краткая форма АПС с кодом первого значения глагола *sollen*

Оригинальное предложение в корпусе ЛГИС	Лемма и значение исследуемого слова в словаре	Перевод предложения в корпусе ЛГИС	Перевод исследуемого слова
Dies alles solle er <i>beschwören</i> bei sämtlichen Heiligen, bei der armen Seele seiner Mutter und bei seiner eigenen Ehre. Patrick Süskind. Das Parfum: Die Geschichte eines Mörders (1985)	sollen <sollen-01>	Пусть он <i>поклонится</i> в этом всеми святыми, бедной душой своей матери и собственной честью. Патрик Зюскинд. Парфюмер: история одного убийцы (перевод Э. Венгерова, 1992)	пусть

Проиллюстрируем на примере (табл. 2) используемый принцип формирования связей, разработанный в рамках предлагаемого подхода с использованием методологии лингвистического аннотирования [20, 21].

Таблица 2 содержит АПС в краткой форме для первого значения глагола *sollen* из вышеприведенного примера словарной статьи для этого глагола, сформированное аннотатором¹ и согласованное с экспертами-лексикографами.

В базе данных ЛГИС аннотаторы формируют АПС для конкретного случая употребления исследуемого слова в корпусе (в табл. 2 это глагол *sollen*). Первый столбец АПС содержит оригинальное предложение из корпуса с той или иной словоформой этого слова (в табл. 2 это *solle*, 3-е лицо единственного числа конъюнктива презенса). Во втором столбце указывается лемма (*sollen*) и код (*sollen-01*), который обозначает то значение этого слова в предложении, которое описано в словарной статье под номером 1.

Третий столбец содержит перевод данного предложения, как он представлен в корпусе ЛГИС, и четвертый столбец — перевод исследуемого слова (*пусть*). В примере АПС код *sollen-01* служит гиперссылкой к описанию в словарной БД этого значения слова, т. е. в данном случае *sollen* под номером «1». Таким образом, АПС обеспечивают связи между словарной БД и текстовыми фрагментами в корпусе ЛГИС по значениям ФМЗ-слов и фразем/идиом, так как первый и третий столбцы АПС содержат оригинал и перевод фрагментов, как они представлены в корпусе, а также ссылки на оригинальный текст и его перевод в самом корпусе.

В заключение этого раздела отметим, что для отражения изменений в описаниях значений ФМЗ-слов и фразем/идиом в базах индивидуальных и коллек-

¹ Аннотированное переводное соответствие, краткая форма которого приведена в табл. 2, сформировано А. А. Гончаровым.

тивных знаний хранятся варианты описаний, сформированные по определенным правилам [9, 18, 19].

4 Реализация предлагаемого подхода

Первая версия ЛГИС предназначена для демонстрации решения актуальной, но во многом нерешенной проблемы многоаспектной интеграции многоязычных электронных словарей и параллельных корпусов. Еще одна цель ее создания состоит в том, чтобы дать возможность лексикографам уже в их практической работе обновлять словарные статьи немецко-русского словаря, опираясь на аутентичные примеры из параллельного корпуса, регулярно пополняемого новыми текстовыми данными, что позволит решить проблему устаревания его содержания. Кроме того, в первой версии ЛГИС пользователи словаря могут не только искать нужные словарные статьи, но и непосредственно обращаться из них по гиперссылкам к более широкому контекстному использованию в корпусе рассматриваемых языковых единиц и их переводов.

Создаваемый ресурс предполагает также возможность обращаться к словарю, исходя из корпуса. Иными словами, в случае, когда пользователь ЛГИС нашел в корпусе какое-то слово, смысл которого ему не вполне ясен, он может перейти в словарь и посмотреть словарную статью с описанием этой языковой единицы и возможными ее переводами (если такая статья уже есть в словарной БД ЛГИС).

После завершения работы над первой версией ЛГИС ее словарная БД будет содержать несколько десятков тысяч словарных статей немецко-русского словаря. В корпус ЛГИС уже загружены параллельные немецко-русские тексты общим объемом более 15 млн словоупотреблений, а в базе данных ЛГИС предполагается сформировать около 6000 АПС. Экспериментально уже проверена реализуемость принципа формирования АПС, обеспечивающих связи словарных статей шести немецких модальных глаголов и параллельных текстов общим объемом в 15 млн словоупотреблений (по значениям каждого из шести глаголов). Всего было сформировано около 2000 АПС, в том числе 420 АПС для первого значения глагола *sollen*.

5 Заключение

При использовании существующих подходов к созданию подобных ресурсов другими исследовательскими командами предлагаются решения частных задач интеграции только для одноязычных словарей и корпусов. Предлагаемый новый теоретический подход, который состоит из шести положений, перечисленных в разд. 3, позиционируется как основа решения проблемы многоаспектной интеграции электронных многоязычных словарей с корпусами текстов как по леммам, так и по значениям многозначных слов, фразем и идиом.

Отметим, что информационных лингвистических ресурсов, подобных первой версии ЛГИС, нет ни в России, ни за рубежом. По этой причине основная

цель исследования включает и создание нового подхода к решению проблемы интеграции словарных и корпусных данных, и демонстрацию реализуемости этого подхода с использованием ресурсов, перечисленных в разд. 4.

Литература

1. Geyken A., Wiegand F., Würzner K.-M. On-the-fly generation of dictionary articles for the DWDS website // *Electronic Lexicography in the 21st Century: Conference Proceedings* / Eds. I. Kosem, C. Tiberius. — Brno: Lexical Computing CZ s. r. o., 2017. P. 560–570.
2. Digitales Wörterbuch der deutschen Sprache. <https://www.dwds.de>.
3. Klein W., Geyken A. Das Digitale Wörterbuch der Deutschen Sprache (DWDS) // *Lexicographica*, 2010. Vol. 26. No. 2010. P. 79–96. doi: 10.1515/9783110223231.1.79.
4. Didakowski J., Lemnitzer L., Geyken A. Automatic example sentence extraction for a contemporary German dictionary // *15th EURALEX Congress (International) Proceedings*. — Oslo, Norway: University of Oslo, 2012. P. 343–349.
5. Pölitz C., Bartz T., Morik K., Storrer A. Investigation of word senses over time using Linguistic Corpora // *Text, speech, and dialogue* / Eds. P. Král, V. Matoušek. — Lecture notes in computer science ser. — Cham: Springer, 2015. Vol. 9302. P. 191–198. doi: 10.1007/978-3-319-24033-6.22.
6. Кружков М. Г. Информационные ресурсы контрастивных лингвистических исследований: электронные корпуса текстов // *Системы и средства информатики*, 2015. Т. 25. № 2. С. 140–159. doi: 10.14357/08696527150209. EDN: UBFBRF.
7. 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. doi: 10.14357/19922264160110. EDN: VXDWPP.
8. 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. doi: 10.14357/08696527180415. EDN: VNHOSJ.
9. Гончаров А. А., Зацман И. М., Кружков М. Г. Эволюция классификаций в надкорпусных базах данных // *Информатика и её применения*, 2020. Т. 14. Вып. 4. С. 108–116. doi: 10.14357/19922264200415. EDN: GKWBZT.
10. Кружков М. Г. Концепция построения надкорпусных баз данных // *Системы и средства информатики*, 2021. Т. 31. № 3. С. 101–112. doi: 10.14357/08696527210309. EDN: UMWNIU.
11. Добровольский Д. О. Структура словаря // *Новый большой немецко-русский словарь* / Под ред. Д. О. Добровольского. — М.: АСТ, 2008. Т. 1. С. 7–30.
12. Гончаров А. А., Зацман И. М., Кружков М. Г. Темпоральные данные в лексикографических базах знаний // *Информатика и её применения*, 2019. Т. 13. Вып. 4. С. 90–96. doi: 10.14357/19922264190415. EDN: SDPAUH.
13. Добровольский Д. О., Кретов А. А., Шаров С. А. Корпус параллельных текстов // *Научно-техническая информация. Сер. 2: Информационные процессы и системы*, 2005. № 6. С. 27–42. EDN: PZQPZX.

14. Добровольский Д. О. Параллельный корпус в исследовании конструкций (проблемы эквивалентности и ее словарного описания) // Контенсивные аспекты языка: константность и вариативность. — М.: Флинта, 2016. С. 75–86. EDN: YGYWPN.
15. Добровольский Д. О. Корпусный подход к исследованию фразеологии: новые результаты по данным параллельных корпусов // Вестник Санкт-Петербургского университета. Язык и литература, 2020. Т. 17. Вып. 3. С. 398–411. doi: 10.21638/spbu09.2020.303. EDN: QZIAAB.
16. Zatsman I. Building digital spiral models of knowledge generation // 19th Forum (International) on Knowledge Asset Dynamics Proceedings. — Matera, Italy: Arts for Business Institute, 2024. P. 2185–2196.
17. Добровольский Д. О., Зацман И. М. Модель извлечения знания из параллельных текстов лексикографической информационной системы // Информатика и её применения, 2024. Т. 18. Вып. 3. С. 97–105. doi: 10.14357/19922264240312. EDN: NDNPCF.
18. Гончаров А. А., Зацман И. М., Кружков М. Г., Лощилова Е. Ю. Отражение эволюции лексикографических знаний в динамических классификационных системах // Информатика и её применения, 2021. Т. 15. Вып. 4. С. 41–49. doi: 10.14357/19922264210406. EDN: MGORMY.
19. Гончаров А. А., Зацман И. М., Кружков М. Г. Представление новых лексикографических знаний в динамических классификационных системах // Информатика и её применения, 2021. Т. 15. Вып. 1. С. 86–93. doi: 10.14357/19922264210112. EDN: OPEFXW.
20. Handbook of linguistic annotation / Eds. N. Ide, J. Pustejovsky. — Dordrecht, The Netherlands: Springer Science + Business Media, 2017. 1568 p. doi: 10.1007/978-94-024-0881-2.
21. Гончаров А. А., Инькова О. Ю., Кружков М. Г. Методология аннотирования в надкорпусных базах данных // Системы и средства информатики, 2019. Т. 29. № 2. С. 148–160. doi: 10.14357/08696527190213. EDN: GNDCJE.

Поступила в редакцию 14.10.2024

Принята к публикации 15.02.2025

INTEGRATION OF A DIGITAL DICTIONARY WITH PARALLEL CORPUS TEXTS: A NEW THEORETICAL APPROACH

D. O. Dobrovol'skiy^{1,2,3} and I. M. Zatsman³

¹Vinogradov Russian Language Institute of the Russian Academy of Sciences, 18/2 Volkhonka Str., Moscow 119019, Russian Federation

²Institute of Linguistics of the Russian Academy of Sciences, 1 bld. 1 Bolshoy Kislovsky Lane, Moscow 125009, Russian Federation

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

Abstract: The present paper considers issues of integrating a digital multilingual dictionary (using the example of a German–Russian dictionary) with texts of

a parallel corpus within the framework of a lexicographic information system that includes three components: (i) a digital multilingual dictionary; (ii) a corpus as a repository of parallel texts; and (iii) a database of annotated translation correspondences and two knowledge bases. The proposed approach to integration is a synthesis of a number of conceptual procedures, including application of the multilevel structuring principle of dictionary entries, formation of annotated translation correspondences for polysemous words and set phrases along with their translations, and providing links between the digital multilingual dictionary and the repository of parallel texts based on individual meanings of polysemous words and set phrases. Until now, such lexicographic information systems have been developed exclusively for monolingual dictionaries with connecting links by lemmas only. The aim of the paper is to describe the proposed approach to integrating a digital multilingual dictionary with texts of a parallel corpus as a theoretical basis for developing a lexicographic information system.

Keywords: lexicographic information system; parallel texts; digital multilingual dictionary; corpus; database of annotated translation correspondences

DOI: 10.14357/08696527250106

EDN: XWJHVC

Acknowledgments

The reported study was funded by the Russian Science Foundation, project No. 24-18-00155, <https://rscf.ru/project/24-18-00155>.

References

1. Geyken, A., F. Wiegand, and K.-M. Würzner. 2017. On-the-fly generation of dictionary articles for the DWDS website. *Electronic Lexicography in the 21st Century: Conference Proceedings*. Eds. I. Kosem and C. Tiberius. Brno: Lexical Computing CZ s. r. o. 560–570.
2. Digitales Wörterbuch der deutschen Sprache. Available at: <https://www.dwds.de> (accessed March 11, 2025).
3. Klein, W., and A. Geyken. 2010. Das digitale Wörterbuch der Deutschen Sprache (DWDS). *Lexicographica* 26(2010):79–96. doi: 10.1515/9783110223231.1.79.
4. Didakowski, J., L. Lemnitzer, and A. Geyken. 2012. Automatic example sentence extraction for a contemporary German dictionary. *15th EURALEX Congress (International) Proceedings*. Oslo, Norway: University of Oslo. 343–349.
5. Pölitiz, C., T. Bartz, K. Morik, and A. Storrer. 2015. Investigation of word senses over time using linguistic corpora. *Text, speech, and dialogue*. Eds. P. Král and V. Matoušek. Lecture notes in computer science ser. Cham: Springer. 9302:191–198. doi: 10.1007/978-3-319-24033-6.22.
6. Kruzhkov, M. G. 2015. Informatsionnye resursy kontrastivnykh lingvisticheskikh issledovaniy: elektronnye korpusa tekstov [Information resources for contrastive studies: Electronic text corpora]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 25(2):140–159. doi: 10.14357/08696527150209. EDN: UBFBRF.
7. Zatsman, I. M., O. Yu. Inkova, M. G. Kruzhkov, and N. A. Popkova. 2016. Predstavlenie krossyazykovykh znaniy o konnektorakh v nadkorporusnykh bazakh daniykh [Representation of cross-lingual knowledge about connectors in supracorpora

- databases]. *Informatika i ee Primeneniya — Inform. Appl.* 10(1):106–118. doi: 10.14357/19922264160110. EDN: VXDWPP.
8. Zatsman, I. M., and M. G. Kruzhkov. 2018. Nadkorpornaya baza dannykh konnektirov: 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. doi: 10.14357/08696527180415. EDN: VNHOSJ.
 9. Goncharov, A. A., I. M. Zatsman, and M. G. Kruzhkov. 2020. Evolyutsiya klassifikatsiy v nadkorpornykh bazakh dannykh [Evolution of classifications in supracorpora databases]. *Informatika i ee Primeneniya — Inform. Appl.* 14(4):108–116. doi: 10.14357/19922264200415. EDN: GKWBZT.
 10. Kruzhkov, M. 2021. Kontseptsiya postroeniya nadkorpornykh baz dannykh [Conceptual framework for supracorpora databases]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 31(3):101–112. doi: 10.14357/08696527210309. EDN: UMWNIU.
 11. Dobrovol'skiy, D. O. 2008. Struktura slovarya [Dictionary structure]. *Novyy bol'shoy nemetsko-russkiy slovar'*. [New large German–Russian dictionary]. Ed. D. O. Dobrovol'skiy. Moscow: AST. 1:7–30.
 12. Goncharov, A. A., I. M. Zatsman, and M. G. Kruzhkov. 2019. Temporal'nye dannye v leksikograficheskikh bazakh znaniy [Temporal data in lexicographic databases]. *Informatika i ee Primeneniya — Inform. Appl.* 13(4):90–96. doi: 10.14357/19922264190415. EDN: SDPAUH.
 13. Dobrovol'skiy, D. O., A. A. Kretov, and S. A. Sharov. 2005. Korpus parallel'nykh tekстов [Corpus of parallel texts]. *Nauchno-tekhnicheskaya informatsiya. Ser. 2: Informatsionnye protsessy i sistemy* [Scientific and Technical Information. Ser. 2: Information processes and systems] 6:27–42. EDN: PZQPZX.
 14. Dobrovol'skiy, D. O. 2016. Parallel'nyy korpus v issledovanii konstruktsiy (problemy ekvivalentnosti i ee slovnogo opisaniya) [Parallel corpus in the study of constructions (problems of equivalence and its dictionary description)]. *Kontensivnye aspekty yazyka: konstantnost' i variativnost'* [Contensive aspects of language: Constancy and variability]. Moscow: Flinta. 75–86. EDN: YGYWPN.
 15. Dobrovol'skiy, D. O. 2020. Korpusnyy podkhod k issledovaniyu frazeologii: novye rezul'taty po dannym parallel'nykh korpusov [Corpus-based approach to phraseology research: New evidence from parallel corpora]. *Vestnik Sankt-Peterburgskogo universiteta. Yazyk i literatura* [Vestnik of Saint Petersburg University. Language and Literature] 17(3):398–411. doi: 10.21638/spbu09.2020.303. EDN: QZIAAB.
 16. Zatsman, I. 2024. Building digital spiral models of knowledge generation. *19th Forum (International) on Knowledge Asset Dynamics Proceedings*. Matera, Italy: Arts for Business Institute. 2185–2196.
 17. Dobrovol'skiy, D. O., and I. M. Zatsman. 2024. Model' izvlecheniya znaniya iz parallel'nykh tekстов leksikograficheskoy informatsionnoy sistemy [A model for extracting knowledge from parallel texts of a lexicographic information system]. *Informatika i ee Primeneniya — Inform. Appl.* 18(3):97–105. doi: 10.14357/19922264240312. EDN: NDNPCF.
 18. Goncharov, A. A., I. M. Zatsman, M. G. Kruzhkov, and E. Yu. Loshchilova. 2021. Ot-razhenie evolyutsii leksikograficheskikh znaniy v dinamicheskikh klassifikatsionnykh

- sistemakh [Capturing evolution of lexicographic knowledge in dynamic classification systems]. *Informatika i ee Primeneniya — Inform. Appl.* 15(4):41–49. doi: 10.14357/19922264210406. EDN: MGORMY.
19. Goncharov, A. A., I. M. Zatsman, and M. G. Kruzhkov. 2021. Predstavlenie novykh leksikograficheskikh znaniy v dinamicheskikh klassifikatsionnykh sistemakh [Representation of new lexicographical knowledge in dynamic classification systems]. *Informatika i ee Primeneniya — Inform. Appl.* 15(1):86–93. doi: 10.14357/19922264210112. EDN: OPEFXW.
20. Ide, N., and J. Pustejovsky, eds. 2017. *Handbook of linguistic annotation*. Dordrecht, The Netherlands: Springer Science + Business Media. 1568 p. doi: 10.1007/978-94-024-0881-2.
21. Goncharov, A. A., O. Yu. Inkova, and M. Kruzhkov. 2019. Metodologiya annotirovaniya v nadkorpusnykh bazakh dannykh [Annotation methodology of supracorpora databases]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(2):148–160. doi: 10.14357/08696527190213. EDN: GNDCJE.

Received October 14, 2024

Accepted February 15, 2025

Contributors

Dobrovol'skij Dmitrij O. (b. 1953) — Doctor of Science in philology, principal scientist, Vinogradov Russian Language Institute of the Russian Academy of Sciences, 18/2 Volkhonka Str., Moscow 119019, Russian Federation; principal scientist, Institute of Linguistics of the Russian Academy of Sciences, 1 bld. 1 Bolshoy Kislovsky Lane, Moscow 125009, Russian Federation; principal scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; dm-dbrv@yandex.ru

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

ГИБРИДНАЯ ИНТЕЛЛЕКТУАЛЬНАЯ МНОГОАГЕНТНАЯ СИСТЕМА КАК МОДЕЛЬ МАЛОГО КОЛЛЕКТИВА СПЕЦИАЛИСТОВ ДЛЯ РЕШЕНИЯ ПРАКТИЧЕСКИХ ПРОБЛЕМ

И. А. Кириков¹, С. В. Листонад², С. Б. Румовская³

Аннотация: Статья посвящена вопросам применения систем искусственного интеллекта (ИИ) для решения практических проблем, характеризующихся неоднородностью, неопределенностью, нечеткостью и динамичностью. Обосновывается применение методов децентрализованного ИИ, позволяющих имитировать работу малых коллективов специалистов, которые традиционно используются для решения подобных проблем. В качестве одного из возможных вариантов реализации данного подхода предлагается концепция гибридных интеллектуальных многоагентных систем (ГиИМАС), в которых интеллектуальные агенты с развитыми моделями предметной области взаимодействуют друг с другом, вырабатывают правила совместного решения проблемы, обмениваются знаниями, необходимыми для ее решения, корректируя собственные представления о ней, распределяют между собой ее части и синтезируют общее решение. Это позволяет интеллектуальной системе проявлять высокую адаптивность, вырабатывая при решении очередной проблемы новый релевантный ей метод. Предложены методология построения, типовая модель и функциональная структура ГиИМАС, выполнена оценка существующих проблем и возможных перспектив развития исследований по данному направлению.

Ключевые слова: коллектив специалистов; решение проблемы; гибридная интеллектуальная многоагентная система; методология построения; концептуальная модель; функциональная структура; многоагентная система; гибридная интеллектуальная система

DOI: 10.14357/08696527250107

EDN: EONKLB

1 Введение

Развитие и усложнение современных систем управления, переход от субъектно-объектной к субъект-метасубъектной парадигме управления [1] порождают практические проблемы, которые требуют не только междисциплинарных знаний, но и способности действовать в условиях неопределенности, нечеткости

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

²Федеральный исследовательский центр «Информатика и управление» Российской академии наук, ser-list-post@yandex.ru

³Федеральный исследовательский центр «Информатика и управление» Российской академии наук, sophiyabr@gmail.com

и динамичности [2–4]. Традиционные методы ИИ демонстрируют ограниченную эффективность в таких сценариях, где ключевыми становятся креативность, адаптивность и синтез разнородных знаний [5]. Методы же генеративного ИИ могут «галлюцинировать», проявлять нестабильность и выдавать результаты, нереализуемые на практике [5, 6]. Таким образом, практические проблемы продолжают решаться традиционным способом с привлечением малых коллективов, которые анализируют проблему с нескольких точек зрения, привлекая специалистов разных направлений. Существенная особенность решения проблем малым коллективом специалистов — потенциальная возможность возникновения синергетического эффекта, когда благодаря взаимодействию участников могут быть получены такие решения проблем или решены такие проблемы, которые не способны получить или решить те же участники, работая порознь [7]. В этой связи малые коллективы специалистов — важный объект исследования, выступающий моделью для создания интеллектуальных систем, которые способны решать практические проблемы без их существенного упрощения.

Для моделирования малых коллективов специалистов, решающих проблемы за круглым столом, в [8] предложен подход на основе ГиИМАС, который интегрирует технологию гибридных интеллектуальных систем [9], обеспечивающую учет неоднородности решаемой проблемы, и аппарат многоагентных систем [10], позволяющий моделировать макроуровневые процессы в коллективах путем организации взаимодействия множества интеллектуальных агентов. Такие агенты — относительно автономные программные сущности, обладающие развитыми моделями предметной области и целеполагания. Таким образом, ГиИМАС — системы третьей (текущей) волны развития ИИ по классификации [11], которые способны обучаться и рассуждать, адаптироваться к новым проблемам, используя накопленный опыт. Исследование и моделирование ГиИМАС механизмов, обуславливающих синергетический эффект в малых коллективах специалистов, позволит разработать методы автоматической интеграции автономных моделей ИИ, усиливающие их возможности по решению практических проблем.

2 Особенности решения проблем в малых коллективах

Малый коллектив специалистов — организованная группа специалистов (обычно до 10 чел.) с общими ценностями, целями и задачами, межличностные отношения которых опосредуются общественно ценным и личностно значимым содержанием совместной деятельности [12]. Группа — совокупность людей, связанных общими чертами, проявляющимися в их совместной деятельности и общении [12]. Коллектив — форма существования группы, ее высшая стадия развития, когда она обретает черты субъекта деятельности [13].

Примерами малых коллективов специалистов, решающих практические проблемы, могут служить: комиссии по ликвидации чрезвычайных ситуаций или аварий; консилиумы врачей, выполняющие обследование, диагностику, прогнозирование исходов заболеваний и формирование планов лечения пациентов

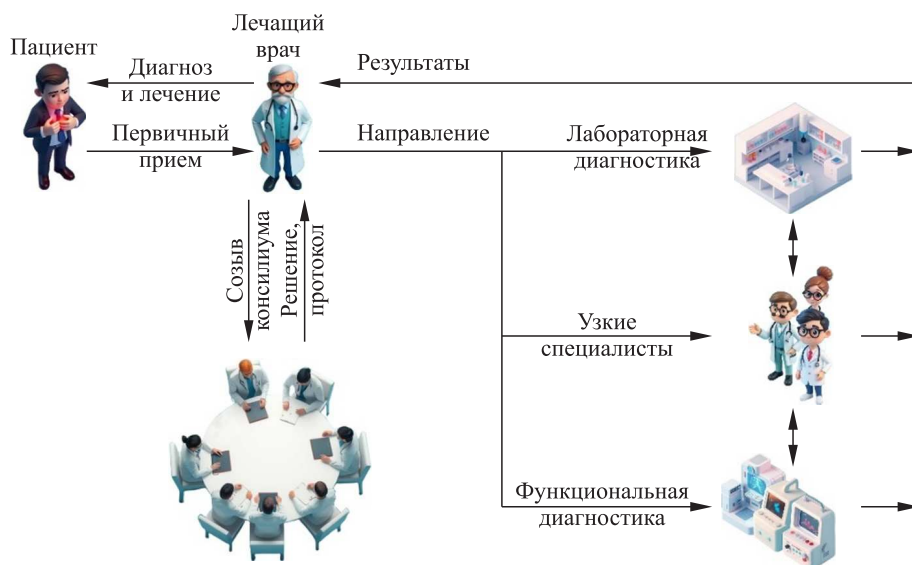


Рис. 1 Схематичное представление процесса диагностики с привлечением консилиума врачей в многопрофильном стационарном лечебном учреждении

со сложными диагнозами, атипичными синдромами и многофункциональными проблемами; коллективы руководителей машиностроительных или транспортно-логистических предприятий, координирующие в ходе планерок работу служб и подразделений для достижения поставленных целей развития организации, и др. [9, 14].

На рис. 1 схематично представлен процесс диагностики с привлечением консилиума в многопрофильном стационарном лечебном учреждении [14]. При поступлении пациент направляется к лечащему врачу в соответствии с профилем основного заболевания, например кардиологу при подозрении на сердечно-сосудистые патологии. В простых случаях врач может сразу поставить диагноз и начать лечение, в более сложных пациент направляется на дополнительное обследование к врачам узких специальностей (невролог, эндокринолог, сосудистый хирург и др.), лабораторную и функциональную диагностику [14], после чего лечащий врач анализирует его результаты, ставит диагноз и назначает лечение. На любом этапе обследования и лечения врач может созвать консилиум в медицинской организации либо вне ее, в том числе дистанционно, при наличии показаний: сложных клинических случаях, нетипичного течения заболевания и др. [15]. Консилиум — совещание нескольких врачей для постановки диагноза, определения прогноза, тактики обследования и лечения [16]. Его работа начинается с доклада лечащего врача о состоянии пациента, данных обследования, использованных методах лечения, после чего проводится осмотр больного [17]. Последующий

обмен мнениями идет в форме свободной дискуссии [17], что позволяет создать интегрированный «метод» диагностики и назначения лечения, сочетающий знания разных областей медицины и релевантный конкретному пациенту. Решение консилиума оформляется протоколом и вносится в медицинскую документацию пациента [16].

Рассмотрение проблем малым коллективом специалистов позволяет уточнить формулировку проблемы в процессе обсуждения, взглянуть на проблему с разных точек зрения, учесть интересы всех сторон, а также потенциально достичь синергетического эффекта в виде неочевидных решений высокого качества, полученных в результате комбинирования знаний и подходов к решению проблем разных участников [7, 18]. Однако рассмотрение проблем в таких коллективах требует больше времени и трудозатрат, чем выработка индивидуальных решений, а итоговые компромиссные решения могут никого не устроить [18].

В отличие от малых коллективов, решение проблем в больших группах характеризуется сложностью визуального и устного общения, ограниченными возможностями для каждого члена высказать свое мнение и донести свою точку зрения, а также более выраженным проявлением эффекта Рингельмана (социальной лености), выражающемся в снижении активности участников и их вклада в обсуждение проблемы [19]. Верхняя граница численности малого коллектива, позволяющая отличить его от большой группы, четко не очерчена, рекомендуется принимать ее в соответствии с потребностью совместной групповой деятельности по решению проблем, для которых он создается [13]. Традиционно считается, что она соответствует «магическому числу» Дж. Миллера 7 ± 2 , но можно встретить оценки до 20–40 чел. При численности более 10 чел. возрастает число подгрупп в коллективе, вероятность противостояния лицу, принимающему решения (ЛПР), а также осложняется координация работы участников [20].

В [21] рассмотрены 12 основных коллективных эффектов, возникающих в малых коллективах специалистов. Отмечается основополагающая роль самоорганизации для возникновения других эффектов положительной направленности, в частности синергии [9]. Исследование механизмов возникновения, развития и затухания коллективных эффектов в ГиИМАС, а также их влияния на конечные решения системы, позволит разработать механизмы непрямого управления процессами самоорганизации с целью формирования коллектива агентов с заданными характеристиками по эффективности решения проблем, сопоставимыми с возможностями реальных коллективов специалистов (при наличии в системе агентов, способных индивидуально решать проблемы на уровне таких специалистов). Другая цель компьютерного моделирования коллективных эффектов с использованием ГиИМАС — исследование их природы, выявление, описание и подтверждение роли в них универсальных закономерностей, не зависящих от природы «интеллекта» и степени влияния человеческого фактора. Это позволит повысить эффективность работы реальных коллективов, а также заложит основы для создания гетерогенных коллективов, в которых решения принимаются в результате диалога специалистов и искусственных агентов.

3 Основные стадии процесса проектирования гибридных интеллектуальных многоагентных систем

В жизненном цикле проектирования ГиИМАС для решения практических проблем путем моделирования малых коллективов специалистов выделяются шесть стадий, которые в идеальном случае выполняются последовательно, но на практике каждая стадия предполагает возможный возврат на любую из предыдущих при необходимости устранения допущенных на ней недочетов:

- (1) на стадии концептуального описания системные аналитики формулируют проблему, предварительно описывая ее на естественном языке в ходе интервью с ЛПР и активными группами [22]. Затем описывается комплекс проблем, связанных с рассматриваемой, чтобы учесть интересы всех заинтересованных лиц (стейкхолдеров), для чего может применяться технология Дж. Уорфилда, метафоры организации, взгляды стейкхолдеров с разных точек зрения и др. [23, 24]. Описание комплекса проблем — основа для формирования дерева или сети целей проектирования ГиИМАС, формируемых в ходе собеседований с каждым стейкхолдером. Стадия завершается выбором критериев и ограничений, которые в совокупности должны быть релевантной количественной моделью построенного дерева или сети целей;
- (2) на стадии моделирования выполняется формализованное описание проблемы, решающего ее коллектива специалистов, если он существует, и проектируемой ГиИМАС. Проблема описывается на макро- и микроуровне. Макроуровневая модель представляет проблему как «черный ящик», включая ее место в метапроблеме, цели, критерии и исходные данные. Микроуровневая модель раскрывает состав и структуру проблемы, ее подпроблемы и связи между ними, специфицируя цели, критерии, исходные данные и методы решения для каждой из них. Модель коллектива отражает ситуацию решения проблемы «как есть сейчас», фиксирует текущий состав специалистов, их роли, задачи, знания и методы, а также порядок взаимодействия. Модель коллектива выступает основой при проектировании ГиИМАС и оценке возможных альтернативных конфигураций. При моделировании ГиИМАС специфицируется множество агентов, протоколы взаимодействия, языки передачи сообщений, базовая онтология, модель окружающей среды, архитектуры и макроуровневые эффекты;
- (3) стадия разработки архитектуры ГиИМАС включает разработку ее функциональной структуры и архитектур ее агентов в соответствии с моделью, построенной на стадии проектирования. Функциональная структура представляется схемами-диаграммами, определяющими подсистемы ГиИМАС, распределение агентов по ним, функционал агентов, допустимые языки и протоколы взаимодействия, технологические элементы, потоки информации и управления, отношения между агентами. Разра-

ботка архитектур агентов выполняется для тех ролей агентов из функциональной структуры, которым не релевантны агенты, ранее реализованные для других систем. Архитектура агента — схема, описывающая состав, структуру и взаимосвязь его функций-блоков. Для каждой функции-блока указывается существующий метод или алгоритм реализации либо отмечается необходимость его разработки на следующей стадии;

- (4) стадия разработки технического проекта ГиИМАС обеспечивает построение методов, алгоритмов, баз данных, онтологий и других элементов, необходимость создания которых отмечена на предыдущей стадии. Порядок их разработки регламентируется соответствующей методологией создания каждого из этих элементов и здесь не рассматривается. Стадия завершается оформлением технического проекта, предписывающего для каждого блока каждого агента и системы в целом их состав и структуру, представление данных, алгоритм исполнения, спецификацию технических средств [25];
- (5) стадия программной реализации и отладки включает разработку программного кода ГиИМАС и его тестирование. Рекомендуется программную реализацию выполнять на платформе JaCaMo [26] с применением Java для программирования отдельных алгоритмов и методов, а также настройки платформы [27]. В ходе тестирования отлаживаются отдельные функции, агенты и технологические элементы, а также система в целом;
- (6) стадия оценки эффективности ГиИМАС обеспечивает сбор информации о работе системы и ее сопоставление с установленными целями в лабораторных условиях, в рамках тестовой эксплуатации и сопровождения после внедрения. В лабораторных условиях специалисты проводят анализ в виртуальной среде, используя для оценки цели, критерии и ограничения, сформированные на стадии концептуального описания проблемы. При тестовой эксплуатации оценка работы системы проводится ее пользователями и стейкхолдерами путем сравнения эффективности ГиИМАС и традиционного метода решения поставленных проблем в ходе их параллельного применения. Сопровождение после внедрения позволяет собирать жалобы и предложения в процессе эксплуатации.

В результате исследований по разработке ГиИМАС для моделирования различных аспектов деятельности малых коллективов специалистов, в частности самоорганизации на основе анализа целей агентов [9], процессов гетерогенного коллективного мышления [28], управления конфликтами агентов [29], формирования сплоченного коллектива [30], разработаны типовая модель и функциональная структура ГиИМАС. Их применение позволит снизить трудозатраты на стадиях моделирования и построения архитектур при проектировании индустриальных образцов ГиИМАС, а также при исследовании других групповых эффектов в малых коллективах специалистов.

4 Типовая модель гибридной интеллектуальной многоагентной системы

Типовая модель ГиИМАС описывается кортежем [29]

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

где AG — множество агентов, описываемое выражением (2); env — концептуальная модель внешней среды системы; INT — множество элементов структурирования взаимодействий агентов (3); ORG — множество архитектур ГиИМАС, соответствующих конкретным функциям (ролям) агентов и установившимся отношениям между ними; MLP — множество концептуальных моделей макроуровневых процессов в ГиИМАС, в зависимости от учитываемых системой коллективных эффектов содержащее модели самоорганизации на основе анализа целей агентов [9], гетерогенного коллективного мышления [28], процесса управления конфликтами агентов [29], процессов согласования целей, онтологий и протоколов, обеспечивающих моделирование сплоченности агентов [30].

Множество агентов из формулы (1) представляется выражением

$$\text{AG} = \{ \text{ag}^{\text{dm}}, \text{ag}^{\text{fc}}, \text{ag}^{\text{med}}, \text{ag}^{\text{int}}, \text{ag}^{\text{pc}} \} \cup \text{AG}^{\text{sp}} \cup \text{AG}^{\text{tr}} \cup \text{AG}^{\text{it}}, \quad (2)$$

где ag^{dm} — агент, принимающий решения (АПР), который моделирует рассуждения ЛПР по распределению подпроблем между агентами-специалистами (АС) и интеграции полученных ими частных решений; ag^{fc} — агент-фасилитатор (АФ), отвечающий за организацию эффективного взаимодействия между агентами системы и опосредованное управление макроуровневыми процессами; ag^{med} — агент-посредник, обеспечивающий поиск агентов с требуемыми функциями для решения частей проблемы; ag^{int} — интерфейсный агент, который предоставляет пользователю возможности по взаимодействию с системой; ag^{pc} — агент контроля протокола, отслеживающий соблюдение агентами согласованных проблемно-ориентированных протоколов; AG^{sp} — подмножество АС, имитирующих работу реальных специалистов по решению поставленной проблемы; AG^{tr} — подмножество агентов-переводчиков, которые обеспечивают взаимодействие пар агентов i, j , не владеющих каким-либо единым языком передачи сообщений: $\text{LANG}_i^{\text{ag}} \cap \text{LANG}_j^{\text{ag}} = \emptyset$ (см. (4)); AG^{it} — подмножество агентов интеллектуальных технологий, обеспечивающих гибридную компоненту системы и учет инструментальной неоднородности поставленной проблемы.

Элементы структурирования взаимодействий (1) описываются выражением

$$\text{INT} = \{ \text{prot}_{\text{bsc}}, \text{PRC}, \text{LANG}, \text{ont}_{\text{bsc}} \}, \quad (3)$$

где prot_{bsc} — базовый протокол, обеспечивающий минимально необходимое взаимодействие для совместной выработки агентами проблемно-ориентированных протоколов; PRC — множество примитивов для конструирования проблемно-ориентированных протоколов АС и АПР; LANG — множество языков

передачи сообщений [31, 32]; ont_{bsc} — базовая онтология, обеспечивающая разбор агентами семантики сообщений.

Агент $\text{ag} \in \text{AG}$ из формулы (2) представляется кортежем

$$\text{ag} = \langle \text{id}^{\text{ag}}, \text{gl}^{\text{ag}}, \text{LANG}^{\text{ag}}, \text{ont}^{\text{ag}}, \text{ACT}^{\text{ag}} \rangle, \quad (4)$$

где id^{ag} — идентификатор; gl^{ag} — нечеткая цель; $\text{LANG}^{\text{ag}} \subseteq \text{LANG}$ — множество языков, которыми «владеет» агент; ont^{ag} — модель предметной области (онтология) агента; ACT^{ag} — множество реализуемых агентом действий.

Действие $\text{act}^{\text{ag}} \in \text{ACT}^{\text{ag}}$ описывается кортежем

$$\text{act}^{\text{ag}} = \langle \text{met}_{\text{act}}^{\text{ag}}, \text{it}_{\text{act}}^{\text{ag}} \rangle,$$

где met_{act} — метод реализации действия; it_{act} — интеллектуальная технология, использованная при построении met_{act} .

Таким образом, функция ГиИМАС описывается выражением

$$\text{act}_{\text{himas}} = \left(\bigcup_{\text{ag} \in \text{AG}} \text{ACT}_{\text{ag}} \right) \cup \text{act}_{\text{col}}, \quad \left| \bigcup_{\text{ag} \in \text{AG}} \bigcup_{\text{act} \in \text{ACT}^{\text{ag}}} \text{it}_{\text{act}}^{\text{ag}} \right| \geq 2,$$

где act_{col} — коллективная функция ГиИМАС, которую агенты формируют динамически, взаимодействуя в соответствии с базовым и проблемно-ориентированными протоколами; ограничение означает, что агенты ГиИМАС должны использовать как минимум две интеллектуальные технологии [9].

5 Типовая функциональная структура гибридной интеллектуальной многоагентной системы

Для компьютерного моделирования коллективных эффектов в малом коллективе специалистов при решении практических проблем предлагается типовая функциональная структура ГиИМАС, представленная на рис. 2.

Как видно из рис. 2, в структуре ГиИМАС можно выделить четыре основные части: интерфейсная подсистема; подсистема гибридизации; подсистема решения проблем; подсистема управления. Интерфейсная подсистема содержит одного агента ag^{int} (2), реализующего взаимодействие пользователя с системой. Этот агент обеспечивает ввод исходных данных и вывод результатов решения проблемы, а также визуализирует макроуровневые процессы MLP, применяя, например, методы визуализации конфликта [33] и сплоченности агентов [34].

Подсистема управления включает базовую онтологию ont_{bsc} (3), несколько агентов-переводчиков из подмножества AG^{tr} , агента-посредника ag^{med} , агента контроля протокола ag^{pc} и АФ ag^{fc} (2). Базовая онтология — семантическая сеть, обеспечивающая однозначную трактовку агентами смысла передаваемой информации. Агенты-переводчики переводят сообщения с одного языка межагентного взаимодействия на другой, что может потребоваться если множества языков, которыми владеют взаимодействующие агенты не пересекаются. Агент-посредник — служба «желтых страниц», он хранит имена, модели и возможности

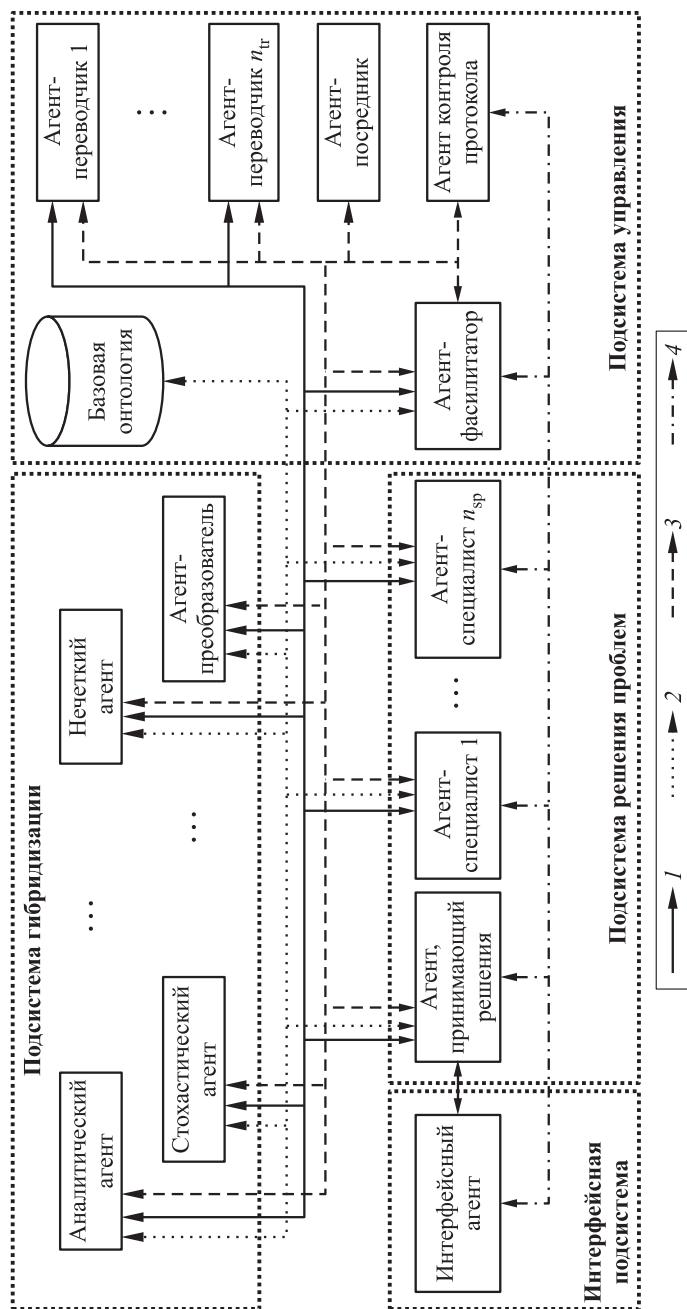


Рис. 2 Функциональная структура ГИИМАС: 1 — отношения агентов: запросы информации, перевода сообщений, помощи в решении задач, результатов их решения; 2 — взаимодействие (получение сведений из модели, обновление модели) агентов с базовой онтологией; 3 — отношения агентов: запросы имен и адресов агентов с заданными возможностями; 4 — отношения фасилитации, инициализирующие процессы согласования целей, моделей предметной области и протокола взаимодействия агентами подсистемы решения проблем

агентов, предоставляя эту информацию любому из агентов системы по запросу, что позволяет им обращаться за помощью в решении частей проблемы к наиболее релевантным агентам. Агент контроля протокола отслеживает соглашения между агентами системы по изменению проблемно-ориентированного протокола и их соблюдение. АФ идентифицирует ситуацию решения проблемы и проявление макроуровневых эффектов MLP (1), стремится усилить одни и подавить другие, чтобы обеспечить эффективную коллективную работу агентов и потенциальную выработку решений более высокого качества.

Подсистема гибридизации содержит подмножество AG^{it} агентов интеллектуальных технологий, которые представляют собой в каком-то смысле библиотеки функций различных методов формализованного представления систем [9] (аналитических, стохастических, нечетких и др.). Агенты этой подсистемы можно назвать рефлекторными, поскольку они не имеют развитой модели предметной области или целеполагания, но именно они обеспечивают ГиИМАС возможность интегрировать модели и алгоритмы различных интеллектуальных технологий, учитывая неотъемлемую инструментальную неоднородность практических проблем. Агенты этой подсистемы выполняют задачи, которые им поручают другие агенты, следуя заложенным алгоритмам, моделям и методам в рамках своей интеллектуальной технологии.

Подсистема решения проблем моделирует работу коллектива специалистов под руководством ЛПР по обсуждению и поиску решения проблемы. Агент, принимающий решения, ag^{dm} (2) моделирует рассуждения ЛПР: распределяет подпроблемы между АС из подмножества AG^{sp} (2), собирает результаты, оценивает достижение требуемого уровня согласованности коллективного решения или другого критерия останова, после чего завершает работу ГиИМАС над проблемой либо инициирует повторное ее рассмотрение. Агенты-специалисты моделируют рассуждения специалистов, вырабатывая решения назначенных им подпроблем или проблемы в целом с учетом собственных целей и моделей предметной области. В общем случае модели предметной области, цели и варианты проблемно-ориентированных протоколов у агентов могут различаться, поэтому для эффективного решения поставленной проблемы агенты согласовывают их между собой по запросу АФ. Таким образом, ГиИМАС способна динамически корректировать свой алгоритм работы, каждый раз подстраиваясь под очередную проблему.

6 Актуальные проблемы и направления развития технологии гибридных интеллектуальных многоагентных систем

Несмотря на широкие возможности ГиИМАС по моделированию коллективного решения проблем в условиях гетерогенности, неопределенности, нечеткости и динамизма, на современном этапе развития данной технологии существует множество ограничений, затрудняющих их широкое практическое применение и требующих проведения дальнейших исследований:

- ограниченные возможности идентификации структуры не встречавшихся ранее слабоформализуемых проблем [35]. Применяемые для этих целей автоматизированные методы требуют участия экспертов, что делает их неприменимыми к динамичным проблемам. Большинство методов автоматической идентификации применимы лишь к формализованным проблемам [36]. Как показано в [36], методы на основе больших языковых моделей, способные работать со слабо формализованными проблемами, даже в лабораторных условиях недостаточно эффективны для решения новых проблем;
- высокая вычислительная сложность моделирования рефлексивных процессов. Под рефлексией понимается способность агента строить и использовать в рассуждениях модели других агентов (учитывая их потенциальную способность к рефлексии), себя самого, своих действий и рассуждений, коллектива агентов в целом как метасубъекта [37, 38]. Рефлексивные способности агентов позволяют сократить длительность переговорных процессов, обеспечить самоорганизацию системы в сильном смысле [39]. Исследование рефлексивно-активных систем искусственных гетерогенных интеллектуальных агентов [38] показало, что рефлексивное моделирование агентами друг друга принципиально возможно, но представляется вычислительно сложным. Существующие работы по созданию рефлексивных агентов на основе больших языковых моделей сосредоточены преимущественно на индивидуальных рефлексивных процессах таких моделей, как централизованных интеллектуальных систем [40, 41]. В этой связи актуальна разработка новых, более эффективных методов моделирования индивидуальной и коллективной рефлексии;
- ограниченность возможностей отдельных агентов по сравнению с реальными специалистами, что затрудняет моделирование коллективных процессов и эффектов при решении практических проблем. Современные искусственные нейронные сети эффективно обрабатывают большие массивы данных, но не способны к глубокому пониманию мира, позволяют выявлять сложные корреляции между входными и выходными признаками, но не формируют причинно-следственные связи [42]. Системы ИИ способны генерировать тексты и изображения, но их «творческие способности» ограничиваются данными, на которых они предварительно обучены [43], в то время как инновации требуют абстрактного мышления, фундаментальной подготовки и интуиции. С учетом высокой сложности и, по всей видимости, нецелесообразности создания систем ИИ, имитирующих или превосходящих человеческие возможности [44, 45], необходим анализ зависимостей между моделируемыми агентами характеристиками и особенностями реальных специалистов, проявлениями макроуровневых эффектов и их влиянием на свойства ГиИМАС в целом, в частности качества предлагаемых решений, требуемых вычислительных ресурсов, устойчивости работы и др. Это позволит подбирать для каждой новой проблемы или класса проблем релевантный состав агентов не

только с точки зрения их «специальности», но и способностей к совместной работе в ГиИМАС;

- недостаточная изученность макроуровневых эффектов в малых коллективах специалистов, что затрудняет их релевантное компьютерное моделирование. Одна из объективных причин, ограничивающих данные исследования, — существенная роль в коллективном взаимодействии неформальной коммуникации и имплицитных норм [46]. При этом на проявление макроуровневых эффектов, например сплоченности коллектива, требуются недели и месяцы его совместной работы [47]. Кроме того, исследования малых коллективов находятся на стыке таких дисциплин, как социология, психология, теория организаций и кибернетика, разрыв между которыми затрудняет интеграцию результатов, полученных каждой из них. В этой связи актуальны междисциплинарные исследования психологов, социологов и инженеров в области информатики и ИИ для создания релевантных моделей макроуровневых эффектов в малых коллективах;
- недостаточная проработка вопросов безопасности. Гибридная интеллектуальная многоагентная система, как любая многоагентная система, обладает уязвимостями на уровне отдельного агента (подделка сообщений, несанкционированный доступ к агенту, атаки агентов и др.) и системы в целом (DoS-атаки, подделка агентов, подделка служб и др.), которые подробно рассмотрены в [48, 49]. Кроме того, интеллектуальные агенты системы, особенно использующие модели машинного обучения в своих рассуждениях, могут быть подвержены атакам через данные для обучения, такими как отравление данных, уклонение, обратная разработка и др. [50–52]. В контексте ГиИМАС эти классы уязвимостей взаимосвязаны: атаки на данные могут привести к системным проблемам, и наоборот. Кроме того, особое внимание должно уделяться рефлексивным процессам в системе, если они моделируются. Эти процессы могут как способствовать повышению безопасности, так и содержать потенциальные уязвимости. Хотя в литературе были представлены разнообразные методы обеспечения безопасности для отдельных типов уязвимостей систем распределенного ИИ, необходимы решения, которые обеспечат комплексный подход к информационной и функциональной безопасности таких систем [49];
- сложность интерпретации человеком сообщений агентов, записанных на стандартных языках передачи сообщений, например KQML [31] или FIPA ACL [32]. Использование более привычных для человека языков при взаимодействии агентов, в частности естественных или визуальных языков [53], языков профессиональной деятельности [9], позволило бы встраивать в процесс обсуждения реальных специалистов, т. е. реализовать концепцию виртуальных гетерогенных коллективов [54], полноценными участниками которых становятся как люди, так и компьютерные модели экспертов. В этом случае множество искусственных интеллектуальных агентов будет выступать

усилителем коллективного человеческого интеллекта или интеллектуальной системой поддержки групповых решений. Хотя методы на основе больших языковых моделей, в частности сократические модели [55], позволяют организовать взаимодействие агентов на естественном языке, возникает проблема корректной интерпретации семантики [56] передаваемых сообщений всеми участниками таких коллективов, а также учета глобальной структуры текста, которая решается современными методами недостаточно качественно [57];

- сложность отладки, а также объяснения полученных решений. В общем случае, если повторно запустить ГиИМАС с теми же параметрами и входными данными, результат может отличаться, что обусловлено недетерминизмом параллельных асинхронных процессов [58], а также вероятностным характером эвристик, которые могут применяться при проектировании отдельных агентов. Для повышения устойчивости выдаваемых решений в [59] предлагается подход с использованием запуска множества копий системы, которые работают параллельно над одной проблемой, после чего выбирается лучшее решение. Такой подход требует значительного увеличения вычислительных ресурсов, при этом не снимает проблему, а лишь маскирует ее проявления.

Кроме перечисленных проблем моделирования коллективного решения проблем средствами ГиИМАС следует отметить ограничения, наследуемые от методологии многоагентных систем, которые затрудняют их широкое промышленное применение, в частности отсутствие четких определений и соглашений по основным понятиям, устаревшие и редко используемые на практике стандарты FIPA [32], отсутствие гибких промышленных методов и технологий для разработки и др. Анализ и пути решения этих проблем рассмотрены в [58].

7 Заключение

В статье обоснована необходимость перехода к моделям, воспроизводящим динамику малых коллективов специалистов, где множество агентов с разными моделями предметной области, целями и стратегиями совместно ищут решения через конфликт идей, распределение ролей и синтез знаний. Такой подход не только расширит спектр решаемых ИИ проблем, но и позволит преодолеть ключевые ограничения, такие как неспособность алгоритмов к адаптации и этические риски, связанные с централизованным принятием решений. Моделирование малых коллективов позволит создавать децентрализованные интеллектуальные системы, способные вырабатывать свой метод решения, релевантный каждой новой проблеме, аналогично тому, как медицинский консилиум подбирает тактику диагностики и лечения под конкретного пациента. С этой целью в работе предложена концепция ГиИМАС, описаны основные стадии ее проектирования, результат которого — программная реализация системы для решения поставленной проблемы, характеризующейся слабой формализацией, неоднородностью и сетевой структурой. Анализ особенностей ГиИМАС, моделирующих различные мак-

роуровневые эффекты в коллективах специалистов, позволил сформулировать также типовую модель и функциональную структуру таких систем. Кроме того, выполнен анализ актуальных проблем и перспектив развития предложенной концепции, который продемонстрировал актуальность дальнейших исследований в данном направлении.

Литература

1. Авдеева З. К., Барышников П. Ю., Журенков Д. А. и др. Стратегическое целеполагание в ситуационных центрах развития / Под ред. В. Е. Лепского, А. Н. Райкова. — М.: Когито-Центр, 2018. 320 с. EDN: YUAMPJ.
2. Спиридонов В. Ф. Психология мышления: Решение задач и проблем. — М.: Генезис, 2006. 319 с. EDN: XJNHLR.
3. Нариньяни А. С. Инженерия знаний и НЕ-факторы: краткий обзор-08 // Вопросы искусственного интеллекта, 2008. Вып. 1. С. 61–77.
4. Malone T. W. Superminds: The surprising power of people and computers thinking together. — New York, NY, USA: Hachette UK, 2018. 384 p.
5. Kulinc H. K., Kecioglu O. F. Generative artificial intelligence: A historical and future perspective // Academic Platform J. Engineering Smart Systems, 2024. Vol. 12. Iss. 2. P. 47–58. doi: 10.21541/apjess.1398155.
6. Егорова А. Ю., Зацман И. М., Романенко В. О. Машинный перевод с помощью ChatGPT: мониторинг воспроизводимости результатов // Системы и средства информатики, 2023. Т. 33. № 3. С. 117–128. doi: 10.14357/08696527230310. EDN: LTOWWN.
7. Larson J. R., Jr. In search of synergy in small group performance. — New York, NY, USA: Psychology Press, 2010. 427 p.
8. Листопад С. В., Кириков И. А. Разрешение конфликтов в гибридных интеллектуальных многоагентных системах // Информатика и её применения, 2022. Т. 16. Вып. 1. С. 54–60. doi: 10.14357/19922264220108. EDN: VHRAQC.
9. Колесников А. В., Кириков И. А., Листопад С. В. Гибридные интеллектуальные системы с самоорганизацией: координация, согласованность, спор. — М.: ИПИ РАН, 2014. 189 с.
10. Тарасов В. Б. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. — М.: Эдиториал УРСС, 2002. 352 с. EDN: ZQQRPP.
11. 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. doi: 10.14357/08696527210105. EDN: IJDSO.
12. Петровский А. В., Шпалинский В. В. Социальная психология коллектива. — М.: Просвещение, 1978. 176 с.
13. Андреева Г. М. Социальная психология. — М.: Аспект пресс, 2009. 363 с.
14. Румовская С. Б., Листопад С. В., Мацула В. Ф., Николаев В. В. Методика автоматизированной диагностики с помощью виртуального консилиума на примере

- артериальной гипертензии // Информатика, управление и системный анализ: Труды V Всеросс. научн. конф. молодых ученых с международным участием. — Ростов-на-Дону: РГЭУ (РИНХ), 2018. С. 83–93. EDN: ХТТИQH.
15. Об организации постоянно действующих консилиумов врачей при оказании медицинской помощи населению города Москвы: Приказ Департамента здравоохранения города Москвы от 7 июня 2017 года № 410 (с изменениями на 13 января 2023 года). <https://docs.cntd.ru/document/456088410>.
 16. Об основах охраны здоровья граждан в Российской Федерации: Федеральный закон № 323-ФЗ от 21.11.2011 (с изменениями на 28 декабря 2024 года). <https://docs.cntd.ru/document/902312609>.
 17. Тихилов Р. М., Кустов В. М. Медицинский (врачебный) консилиум — история и практика // Уральский медицинский ж., 2010. Т. 72. Вып. 7. С. 5–8.
 18. Самсонова М. В., Ефимов В. В. Технология и методы коллективного решения проблем. — Ульяновск: УЛГТУ, 2003. 152 с.
 19. Jacobs G., Farrell T. S. C. Advantages of combining small and large groups: Implications for second language education // Brit. S. Ap. L., 2019. Vol. 7. Iss. 2. P. 61–70. doi: 10.33508/bw.v7i2.2265.
 20. Румовская С. Б., Кириков И. А. Методы моделирования и визуального представления конфликта в малом коллективе экспертов, решающих проблемы (обзор) // Информатика и её применения, 2019. Т. 13. Вып. 3. С. 122–130. doi: 10.14357/19922264190317. EDN: VNSTEE.
 21. Кириков И. А., Колесников А. В., Листопад С. В. Моделирование систем поддержки принятия решений синергетическим искусственным интеллектом // Информатика и её применения, 2013. Т. 7. Вып. 3. С. 62–69. EDN: RFLJLB.
 22. Ларичев О. И. Вербальный анализ решений. — М.: Наука, 2006. 181 с.
 23. Новиков А. М., Новиков Д. А. Методология. — М.: Синтег, 2007. 668 с.
 24. Тарасенко Ф. П. Прикладной системный анализ. — М.: КНОРУС, 2010. 224 с. EDN: QMUKWJ.
 25. Черушева Т. В. Проектирование программного обеспечения. — Пенза: ПГУ, 2014. 172 с.
 26. Boissier O., Bordini R. H., Hübner J. F., Ricci A. Multi-agent oriented programming: Programming multi-agent systems using JaCaMo. — Intelligent robotics and autonomous agents ser. — Cambridge, MA, USA: The MIT Press, 2020. 264 p.
 27. Смирнов С. С., Смольянинова В. А. Введение в разработку многоагентных систем в среде Jason. Основы программирования на языке AgentSpeak. — М.: МИРЭА, 2009. 136 с.
 28. Колесников А. В., Листопад 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. EDN: YPEYCT.
 29. Листопад 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. EDN: JHSVES.

30. Лустонад С. В. Модель сплоченной гибридной интеллектуальной многоагентной системы // Гибридные и синергетические интеллектуальные системы: Мат-лы V Всеросс. Поспеловской конф. с международным участием. — Калининград: БФУ им. И. Канта, 2020. С. 437–444. EDN: UAMUFA.
31. Finin T., Fritzson R., McKay D., McEntire R. KQML as an agent communication language // 3rd Conference (International) on Information and Knowledge Management Proceedings. — New York, NY, USA: ACM, 1994. P. 456–463. doi: 10.1145/191246.191322.
32. FIPA ACL message structure specification. <http://www.fipa.org/specs/fipa00061/index.html>.
33. Румовская С. Б., Кириков И. А. Визуальный язык репрезентации процесса управления конфликтами в гибридных интеллектуальных многоагентных системах // Системы и средства информатики, 2023. Т. 33. № 2. С. 60–70. doi: 10.14357/08696527230206. EDN: JMDWWG.
34. Rumovskaya S. Visualization of team cohesion in hybrid intelligent multi-agent systems // 2nd Conference (International) on Control Systems, Mathematical Modeling, Automation and Energy Efficiency. — IEEE, 2020. P. 620–623. doi: 10.1109/SUMMA50634.2020.9280790.
35. Лустонад С. В., Кириков И. А. Анализ подходов к идентификации структуры проблемы в гибридных интеллектуальных многоагентных системах // Системы и средства информатики, 2024. Т. 34. № 3. С. 35–47. doi: 10.14357/08696527240304. EDN: RSLMXC.
36. Ortiz S. What is Auto-GPT? Everything to know about the next powerful AI tool. <https://www.zdnet.com/article/what-is-auto-gpt-everything-to-know-about-the-next-powerful-ai-tool>.
37. Лефевр В. А. Конфликтующие структуры. — М.: Советское радио, 1973. 158 с.
38. Лустонад С. В. Базовая архитектура рефлексивно-активных систем искусственных гетерогенных интеллектуальных агентов // Информатика и её применения, 2024. Т. 18. Вып. 3. С. 89–96. doi: 10.14357/19922264240311. EDN: UNTQBV.
39. Di Marzo Serugendo G., Gleizes M.-P., Karageorgos A. Self-organization in multi-agent systems // Knowl. Eng. Rev., 2005. Vol. 20. Iss. 2. P. 165–189. doi: 10.1017/S0269888905000494.
40. Introspective agents: Performing tasks with reflection — LlamaIndex. https://docs.llamaindex.ai/en/stable/examples/agent/introspective_agent_toxicity_reduction.
41. Renze M., Guven E. Self-reflection in LLM agents: Effects on problem-solving performance. — Cornell University, 2024. arXiv: 2405.06682 [cs]. 14 p.
42. Marcus G. Deep learning: A critical appraisal. — Cornell University, 2018. arXiv: 1801.00631 [cs]. 27 p.
43. Yadlowsky S., Doshi L., Tripuraneni N. Pretraining data mixtures enable narrow model selection capabilities in transformer models. — Cornell University, 2023. arXiv:2311.00871 [cs]. 13 p.
44. Russell S. J. Human compatible: artificial intelligence and the problem of control. — New York, NY, USA: Viking, 2019. 352 p.

45. *Гринченко С. Н.* Антропогенная «третья» природа: относительно автономный статус ее искусственных интеллектуальных субъектов // Информатика и её применения, 2021. Т. 15. Вып. 4. С. 110–114. doi: 10.14357/19922264210415. EDN: KUYRPD.
46. *Sinz S.* The role of informal social interactions in predicting team cohesion and performance: Master Thesis. — Montreal: Concordia University, 2020. 57 p.
47. *Chiocchio F., Essiembre H.* Cohesion and performance: A meta-analytic review of disparities between project teams, production teams, and service teams // Small Gr. Res., 2009. Vol. 40. Iss. 4. P. 382–420. doi: 10.1177/1046496409335103.
48. *Jung Y., Kim M., Masoumzadeh A., Joshi J. B. D.* A survey of security issue in multi-agent systems // Artif. Intell. Rev., 2012. Vol. 37. Iss. 3. P. 239–260. doi: 10.1007/s10462-011-9228-8.
49. *Hedin Y., Moradian E.* Security in multi-agent systems // Procedia Comput. Sci., 2015. Vol. 60. P. 1604–1612. doi: 10.1177/1046496409335103.
50. *Зацаринный А. А., Сучков А. П.* Некоторые подходы к анализу факторов, влияющих на информационную безопасность систем искусственного интеллекта // Системы и средства информатики, 2023. Т. 33. № 3. С. 95–107. doi: 10.14357/08696527230308. EDN: QNXZBN.
51. *Грушо А. А., Грушо Н. А., Забежайло М. И., Писковский В. О., Тимонина Е. Е., Шоргин С. Я.* Логика обмана в машинном обучении // Информатика и её применения, 2024. Т. 18. Вып. 1. С. 78–83. doi: 10.14357/19922264240111. EDN: SCDSHX.
52. *Сучков А. П.* Жизненный цикл кибератак на уязвимости систем машинного обучения // Системы и средства информатики, 2024. Т. 34. № 1. С. 80–91. doi: 10.14357/08696527240107. EDN: QLGHEH.
53. *Колесников А. В., Листопад С. В.* Алгоритм синтеза функциональных гибридных интеллектуальных систем с гетерогенным визуальным полем // Системы и средства информатики, 2017. Т. 27. № 3. С. 112–125. doi: 10.14357/08696527170310. EDN: ZQLXWH.
54. *Кириков И. А., Колесников А. В., Листопад С. В., Румовская С. Б.* Виртуальные гетерогенные коллективы, поддерживающие принятие решений // Системы и средства информатики, 2015. Т. 25. № 3. С. 126–149. doi: 10.14357/08696527150308. EDN: UMULTH.
55. *Zeng A., Attarian M., Ichter B., Choromanski K., Wong A., Welker S., Tombari F., Purohit A., Ryoo M., Sindhvani V., Lee J., Vanhoucke V., Florence P.* Socratic models: Composing zero-shot multimodal reasoning with language. — Cornell University, 2022. arXiv:2204.00598 [cs]. 30 p.
56. *Смирнов И. В.* Разноуровневая обработка естественного языка для интеллектуального поиска и анализа текстов // Искусственный интеллект и принятие решений, 2023. Вып. 1. С. 90–99. doi: 10.14357/20718594230109. EDN: ALPUWY.
57. *Смирнов И. В., Панов А. И., Чуганская А. А., Суворова М. И., Киселёв Г. А., Курузов И. А., Григорьев О. Г.* Персональный когнитивный ассистент: планирование поведения на основе сценариев деятельности // Информатика и её применения, 2022. Т. 16. Вып. 1. С. 46–53. doi: 10.14357/19922264220107. EDN: RAGFXI.

58. Городецкий В. И., Бухвалов О. Л., Скобелев П. О. Современное состояние и перспективы индустриальных применений многоагентных систем // Управление большими системами, 2017. Вып. 66. С. 94–157. EDN: YZAKOV.
59. Чернышев С. А. Проблемы мультиагентных систем и возможные пути их решения // Вестник Российского нового университета. Сер. Сложные системы: модели, анализ и управление, 2023. Вып. 3. С. 231–241. doi: 10.18137/RNU.V9187.23.03.P.231. EDN: DWJCWP.

Поступила в редакцию 09.01.2025

Принята к публикации 15.02.2025

HYBRID INTELLIGENT MULTIAGENT SYSTEM AS A MODEL OF A SMALL TEAM OF SPECIALISTS FOR SOLVING PRACTICAL PROBLEMS

I. A. Kirikov, S. V. Listopad, and S. B. Rumovskaya

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

Abstract: The paper is devoted to the issues of application of artificial intelligence systems for solving practical problems characterized by heterogeneity, uncertainty, fuzzyness, and dynamism. The application of decentralized artificial intelligence methods that allow simulating the work of small teams of specialists, which are traditionally used to solve such problems, is substantiated. As one of the possible options for implementing this approach, the concept of hybrid intelligent multiagent systems is proposed, in which intelligent agents with extensive domain models interact with each other, develop rules for jointly solving the problem, exchange knowledge necessary for its solution, adjusting their own ideas about it, distribute its parts among themselves, and synthesize a common solution. This allows the intelligent system to demonstrate high adaptability developing a new relevant method when solving the next problem. The paper proposes a methodology for constructing a typical model and a functional structure of a hybrid intelligent multiagent system, an assessment of existing problems, and possible prospects for the development of research in this area.

Keywords: team of specialists; problem solving; hybrid intelligent multiagent system; construction methodology; conceptual model; functional structure; multiagent system; hybrid intelligent system

DOI: 10.14357/08696527250107

EDN: EONKLB

References

1. Avdeeva, Z. K., P. Yu. Baryshnikov, D. A. Zhurenkov, *et al.* 2018. *Strategicheskoe tselepolaganie v situatsionnykh tsentrakh razvitiya* [Strategic goal-setting in situational development centers]. Eds. V. E. Lepskiy and A. N. Raykov. Moscow: Cogito-Centre. 320 p. EDN: YUAMPJ.

2. Spiridonov, V. F. 2006. *Psikhologiya myshleniya: Reshenie zadach i problem* [Psychology of thinking: Problem solving]. Moscow: Genesis. 323 p. EDN: XJNHLP.
3. Narinyani, A. S. 2008. Inzheneriya znaniy i NE-factory: kratkiy obzor-08 [Knowledge engineering and nonfactors: A brief overview-08]. *Voprosy iskusstvennogo intellekta* [Artificial Intelligence Issues] 1:61–77.
4. Malone, T. W. 2018. *Superminds: The surprising power of people and computers thinking together*. New York, NY: Hachette UK. 384 p.
5. Kılınç, H. K., and O. F. Kececioğlu. 2024. Generative artificial intelligence: A historical and future perspective. *Academic Platform J. Engineering Smart Systems* 12(2):47–58. doi: 10.21541/apjess.1398155.
6. Egorova, A. Yu., I. M. Zatsman, and V. O. Romanenko. 2023. Mashinnyy perevod s pomoshch'yu ChatGPT: monitoring vosproizvodimosti rezul'tatov [Machine translation by ChatGPT: Monitoring of outcome reproducibility]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 33(3):117–128. doi: 10.14357/08696527230310. EDN: LTOWWN.
7. Larson, J. R., Jr. 2010. *In search of synergy in small group performance*. New York, NY: Psychology Press. 427 p.
8. Listopad, S. V., and I. A. Kirikov. 2022. Razreshenie konfliktov v gibridnykh intellektual'nykh mnogoagentnykh sistemakh [Conflict resolution in hybrid intelligent multiagent systems]. *Informatika i ee Primeneniya — Inform. Appl.* 16(1):54–60. doi: 10.14357/19922264220108. EDN: VHRAQC.
9. Kolesnikov, A. V., I. A. Kirikov, and S. V. Listopad. 2014. *Gibridnye intellektual'nye sistemy s samoorganizatsiey: koordinatsiya, soglasovannost', spor* [Hybrid intelligent systems with self-organization: Coordination, consistency, and dispute]. Moscow: IPI RAN. 189 p.
10. Tarasov, V. B. 2002. *Ot mnogoagentnykh sistem k intellektual'nym organizatsiyam: filosofiya, psikhologiya, informatika* [From multiagent systems to intelligent organizations: Philosophy, psychology, and informatics]. Moscow: Editorial URSS. 352 p. EDN: ZQQRPP.
11. 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. doi: 10.14357/08696527210105. EDN: IJDSO.
12. Petrovskiy, A. V., and V. V. Shpalinskiy. 1978. *Sotsial'naya psikhologiya kolektiva* [Social psychology of the collective]. Moscow: Prosveshchenie. 176 p.
13. Andreeva, G. M. 2009. *Sotsial'naya psikhologiya* [Social psychology]. Moscow: Aspekt Press. 363 p.
14. Rumovskaya, S. B., S. V. Listopad, V. F. Matsula, and V. V. Nikolaev. 2018. Metodika avtomatizirovannoy diagnostiki s pomoshch'yu virtual'nogo konsiliuma na primere arterial'noy gipertenzii [Methodology of automated diagnostics using a virtual consultation on the example of arterial hypertension]. *Informatika, upravlenie i sistemnyy analiz: Trudy V Sseross. nauchn. konf. molodykh uchenykh s mezhdunarodnym uchastiem* [Computer Science, Management, and Systems Analysis: 5th All-Russian

- Scientific Conference of Young Scientists with International Participation Proceedings]. Rostov-on-Don: RSUE (RINH). 83–93. EDN: XTTIQH.
15. Ob organizatsii postoyanno deystvuyushchikh konsiliumov vrachey pri okazanii meditsinskoj pomoshchi naseleniyu goroda Moskvyy: Prikaz Departamenta zdravookhraneniya goroda Moskvyy ot 7 iyunya 2017 goda No.410 (s izmeneniyami na 13 yanvarya 2023 goda) [On the organization of permanent consultations of doctors when providing medical care to the population of the city of Moscow: Order of the Moscow Department of Healthcare dated June 7, 2017 No. 410 (as amended on January 13, 2023)]. Available at: <https://docs.cntd.ru/document/456088410> (accessed March 4, 2025).
 16. Ob osnovakh okhrany zdorov'ya grazhdan v Rossiyskoy Federatsii: Federal'nyy zakon No.323-FZ ot 21.11.2011 (s izmeneniyami na 28 dekabrya 2024 goda) [On the fundamentals of protecting the health of citizens in the Russian Federation: Federal Law No.323-FZ dated November 21, 2011 (as amended on December 28, 2024)]. Available at: <https://docs.cntd.ru/document/902312609> (accessed March 4, 2025).
 17. Tikhilov, P. M., and V. M. Kustov. 2010. Meditsinskiy (vrachebnyy) konsilium — istoriya i praktika [Medical (physician) consultation — history and practice]. *Ural'skiy meditsinskiy zh.* [Ural Medical J.] 72(7):5–8.
 18. 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.
 19. Jacobs, G., and T. S. C. Farrell. 2019. Advantages of combining small and large groups: Implications for second language education. *Brit. S. Ap. L.* 7(2):61–70. doi: 10.33508/bw.v7i2.2265.
 20. Rumovskaya, S. B., and I. A. Kirikov. 2019. Metody modelirovaniya i vizual'nogo predstavleniya konflikta v malom kollektive ekspertov, reshayushchikh problemy (obzor) [Methods of modeling and visual representation of a conflict in a small collective of experts solving problems (review)]. *Informatika i ee Primeneniya — Inform. Appl.* 13(3):122–130. doi: 10.14357/19922264190317. EDN: VNSTEE.
 21. Kirikov, I. A., A. V. Kolesnikov, and S. V. Listopad. 2013. Modelirovanie sistem podderzhki prinyatiya resheniy sinergeticheskim iskusstvennym intellektom [Decision support systems modeling with synergetic artificial intelligence]. *Informatika i ee Primeneniya — Inform. Appl.* 7(3):62–69. EDN: RFLJLB.
 22. Larichev, O. I. 2006. *Verbal'nyy analiz resheniy* [Verbal analysis of decisions]. Moscow: Nauka. 181 p.
 23. Novikov, A. M., and D. A. Novikov. 2007. *Metodologiya* [Methodology]. Moscow: Sinteg. 668 p.
 24. Tarasenko, F. P. 2010. *Prikladnoy sistemnyy analiz* [Applied systems analysis]. Moscow: Knorus. 224 p. EDN: QMUKWJ.
 25. Cherusheva, T. V. 2014. *Proektirovanie programmnogo obespecheniya* [Software design]. Penza: PGU. 172 p.
 26. Boissier, O., R. H. Bordini, J. Hübner, and A. Ricci. 2020. *Multi-agent oriented programming: Programming multi-agent systems using JaCaMo*. Intelligent robotics and autonomous agents ser. Cambridge, MA: The MIT Press. 264 p.

27. Smirnov, S. S., and V. A. Smol'yaninova. 2009. *Vvedenie v razrabotku mnogoagentnykh sistem v srede Jason. Osnovy programmirovaniya na yazyke AgentSpeak* [Introduction to the development of multiagent systems in the Jason environment. Fundamentals of programming in the AgentSpeak language]. Moscow: MIREA. 136 p.
28. 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. EDN: YPEYCT.
29. Listopad, S. V., and I. A. Kirikov. 2020. Metod identifikatsii konfliktov agentov v gibridnykh 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. EDN: JHSVES.
30. Listopad, S. V. 2020. Model' splochennoy gibridnoy intellektual'noy mnogoagentnoy sistemy [Model of a cohesive hybrid intelligent multiagent system]. *Gibridnye i sinergeticheskie intellektual'nye sistemy: Mat-ly V Sseross. Pospelovskoy konf. s mezhdunarodnym uchastiem* [Hybrid and Synergetic Intelligent Systems: 5th All-Russian Pospelovskaya Conference with International Participation Proceedings]. Kaliningrad: IKBFU. 437–444. EDN: UAMUFA.
31. Finin, T., R. Fritzson, D. McKay, and R. McEntire. 1994. KQML as an agent communication language. *3rd Conference (International) on Information and Knowledge Management Proceedings*. New York, NY: ACM. 456–463. doi: 10.1145/191246.191322.
32. FIPA ACL message structure specification. Available at: <http://www.fipa.org/specs/fipa00061/index.html> (accessed March 4, 2025).
33. Rumovskaya, S. B., and I. A. Kirikov. 2023. Vizual'nyy yazyk reprezentatsii protsessa upravleniya konfliktami v gibridnykh intellektual'nykh mnogoagentnykh sistemakh [A visual language of the representation of the conflict management process in hybrid intelligent multiagent systems]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 33(2):60–70. doi: 10.14357/08696527230206. EDN: JMDWWG.
34. Rumovskaya, S. 2020. Visualization of team cohesion in hybrid intelligent multiagent systems. *2nd Conference (International) on Control Systems, Mathematical Modeling, Automation and Energy Efficiency Proceedings*. IEEE. 620–623. doi: 10.1109/SUMMA50634.2020.9280790.
35. Listopad, S. V., and I. A. Kirikov. 2024. Analiz podkhodov k identifikatsii struktury problemy v gibridnykh intellektual'nykh mnogoagentnykh sistemakh [Analysis of approaches to problem structure identification in hybrid intelligent multiagent systems]. *Sistemy i sredstva informatiki — Systems and Means of Informatics* 34(3):35–47. doi: 10.14357/08696527240304. EDN: RSLMXC.
36. Ortiz, S. 2023. What is Auto-GPT? Everything to know about the next powerful AI tool. Available at: <https://www.zdnet.com/article/what-is-auto-gpt-everything-to-know-about-the-next-powerful-ai-tool> (accessed March 4, 2025).

37. Lefebvre, V. A. 1973. *Konfliktuyushchie struktury* [Conflicting structures]. Moscow: Soviet Radio. 158 p.
38. Listopad, S. V. 2024. Bazovaya arkhitektura refleksivno-aktivnykh sistem iskusstvennykh geterogennykh intellektual'nykh agentov [Basic architecture of reflexive-active systems of artificial heterogeneous intelligent agents]. *Informatika i ee Primeneniya — Inform. Appl.* 18(3):89–96. doi: 10.14357/19922264240311. EDN: UNTQBV.
39. Di Marzo Serugendo, G., M.-P. Gleizes, and A. Karageorgos. 2005. Self-organization in multi-agent systems. *Knowl. Eng. Rev.* 20(2):165–189. doi: 10.1017/S0269888905000494.
40. Introspective agents: Performing tasks with reflection. LlamaIndex. Available at: https://docs.llamaindex.ai/en/stable/examples/agent/introspective_agent_toxicity_reduction (accessed March 4, 2025).
41. Renze, M., and E. Guven. 2024. Self-reflection in LLM agents: Effects on problem-solving performance. Cornell University. 14 p. Available at: <https://arxiv.org/abs/2405.06682> (accessed March 4, 2025).
42. Marcus, G. 2018. Deep learning: A critical appraisal. Cornell University. 27 p. Available at: <https://arxiv.org/abs/1801.00631> (accessed March 4, 2025).
43. Yadlowsky, S., L. Doshi, and N. Tripuraneni. 2023. Pretraining data mixtures enable narrow model selection capabilities in transformer models. Cornell University. 13 p. Available at: <https://arxiv.org/abs/2311.00871> (accessed March 4, 2025).
44. Russell, S. J. 2019. *Human compatible: Artificial intelligence and the problem of control*. New York, NY: Viking. 352 p.
45. Grinchenko, S. N. 2021. Antropogennaya “tret'ya” priroda: otnositel'no avtonomnyy status ee iskusstvennykh intellektual'nykh sub"ektov [Anthropogenic “third” nature: The relatively autonomous status of its artificial intellectual subjects]. *Informatika i ee Primeneniya — Inform. Appl.* 15(4):110–114. doi: 10.14357/19922264210415. EDN: KUYRPD.
46. Sinz, S. 2020. The role of informal social interactions in predicting team cohesion and performance. Montreal: Concordia University. Master Thesis. 57 p.
47. Chiocchio, F., and H. Essiembre. 2009. Cohesion and performance: A meta-analytic review of disparities between project teams, production teams, and service teams. *Small Gr. Res.* 40(4):382–420. doi: 10.1177/1046496409335103.
48. Jung, Y., M. Kim, A. Masoumzadeh, and J. B. D. Joshi. 2012. A survey of security issue in multi-agent systems. *Artif. Intell. Rev.* 37(3):239–260. doi: 10.1007/s10462-011-9228-8.
49. Hedin, Y., and E. Moradian. 2015. Security in multi-agent systems. *Procedia Comput. Sci.* 60:1604–1612. doi: 10.1016/j.procs.2015.08.270.
50. Zatsarinny, A. A., and A. P. Suchkov. 2023. Nekotorye podkhody k analizu faktorov, vliyayushchikh na informatsionnyu bezopasnost' sistem iskusstvennogo intellekta [Some approaches to the analysis of factors affecting the information security of artificial intelligence systems]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 33(3):95–107. doi: 10.14357/08696527230308. EDN: QNXZBN.
51. Grusho, A. A., N. A. Grusho, M. I. Zabezhailo, V. O. Piskovskiy, E. E. Timonina, and S. Ya. Shorgin. 2024. Logika obmana v mashinnom obuchenii [Logic of deception in

- machine learning]. *Informatika i ee Primeneniya — Inform. Appl.* 18(1):78–83. doi: 10.14357/19922264240111. EDN: SCDSHX.
52. Suchkov, A. P. 2024. Zhiznennyy tsikl kiberatak na uyazvimosti sistem mashinnogo obucheniya [The lifecycle of cyberattacks on machine learning system vulnerabilities]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 34(1):80–91. doi: 10.14357/08696527240107. EDN: QLGHEH.
 53. Kolesnikov, A. V., and S. V. Listopad. 2017. Algoritm sinteza funktsional'nykh gi-briidnykh intellektual'nykh sistem s geterogennym vizual'nym polem [Algorithm for synthesis of functional hybrid intelligent systems with a heterogeneous visual field]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 27(3):112–125. doi: 10.14357/08696527170310. EDN: ZQLXWH.
 54. Kirikov, I. A., A. V. Kolesnikov, S. V. Listopad, and S. B. Rumovskaya. 2015. Virtual'nye geterogennye kollektivy, podderzhivayushchie prinyatie resheniy [Virtual heterogeneous collectives for supporting decision-making]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 25(3):126–149. doi: 10.14357/08696527150308. EDN: UMULTH.
 55. Zeng, A., M. Attarian, B. Ichter, K. Choromanski, A. Wong, S. Welker, F. Tombari, A. Purohit, M. Ryoo, V. Sindhwani, J. Lee, V. Vanhoucke, and P. Florence. 2022. Socratic models: Composing zero-shot multimodal reasoning with language. Cornell University. 30 p. Available at: <https://arxiv.org/abs/2204.00598> (accessed March 4, 2025).
 56. Smirnov, I. V. 2023. Raznourovnevaya obrabotka estestvennogo yazyka dlya intellektual'nogo poiska i analiza tekstov [Multilevel language processing for intelligent retrieval and text mining]. *Iskusstvennyy intellekt i prinyatie resheniy* [Artificial Intelligence and Decision Making] 1:90–99. doi: 10.14357/20718594230109. EDN: ALPUWY.
 57. Smirnov, I. V., A. I. Panov, A. A. Chuganskaya, M. I. Suvorova, G. A. Kiselev, I. A. Kuruzov, and O. G. Grigoriev. 2022. Personal'nyy kognitivnyy assistent: planirovanie povedeniya na osnove stsensariiev deyatel'nosti [Personal cognitive assistant: Planning activity with scripts]. *Informatika i ee Primeneniya — Inform. Appl.* 16(1):46–53. doi: 10.14357/19922264220107. EDN: RAGFXI.
 58. Gorodetsky, V. I., O. L. Bukhvalov, and P. O. Skobelev. 2017. Sovremennoe so-stoyanie i perspektivy industrial'nykh primeneniy mnogoagentnykh sistem [Industrial applications of multi-agent systems: Current state and prospects]. *Upravlenie bol'shimi sistemami* [Large-Scale Systems Control] 66:94–157. EDN: YZAKOV.
 59. Chernyshev, S. A. 2023. Problemy mul'tiagentnykh sistem i vozmozhnye puti ikh resheniya [Problems of multi-agent systems and possible ways to solve them]. *Vestnik Rossiyskogo novogo universiteta*. Ser. Slozhnyye sistemy: modeli, analiz i upravlenie [Vestnik of Russian New University. Complex systems: Models, analysis, management ser.] 3:231–241. doi: 10.18137/RNU.V9187.23.03.P.231. EDN: DWJCWP.

Received January 9, 2025

Accepted February 15, 2025

Contributors

Kirikov Igor A. (b. 1955) — Candidate of Science (PhD) in technology, director, Kaliningrad Branch of the Federal Research Center “Computer Science and Control”

of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; baltbipiran@mail.ru

Listopad Sergey V. (b. 1984) — Candidate of Science (PhD) in technology, senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; ser-list-post@yandex.ru

Rumovskaya Sophiya B. (b. 1985) — Candidate of Science (PhD) in technology, senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; sophiyabr@gmail.com

СВОЙСТВА И ОПТИМИЗАЦИЯ САМОСИНХРОННЫХ СХЕМ

*В. Н. Захаров¹, Ю. А. Степченко², Ю. Г. Дьяченко³, Н. В. Морозов⁴,
Л. П. Плеханов⁵, Д. Ю. Степченко⁶*

Аннотация: Статья посвящена краткому обзору тенденций развития и достижений в области самосинхронных (СС) схем. За счет избыточного кодирования данных и двухфазной дисциплины СС-схемы обладают рядом преимуществ в сравнении с синхронными (С) аналогами. Они надежно функционируют при любых задержках элементов, выявляют любые константные неисправности, имеют предельно широкий диапазон работоспособности по напряжению питания и температуре окружающей среды. Практическая разработка СС-устройств разной сложности доказала эффективность СС-решений, особенно в высоконадежных и отказоустойчивых применениях. Статья приводит результаты сравнительных испытаний тестовых образцов С- и СС-схем, оценки быстродействия и сбоеустойчивости СС-схем разного уровня сложности. Описана групповая индикация многоразрядных СС-схем, повышающая их быстродействие на 40% за счет незначительного увеличения (менее 3%) аппаратных затрат. Приведены схемотехнические методы защиты от логических сбоев, улучшающие естественную иммунность СС-схем к неблагоприятным воздействиям и обеспечивающие уровень сбоеустойчивости СС-схем в несколько раз выше, чем у С-аналогов.

Ключевые слова: самосинхронная схема; С-элемент; парафазный; индикация; конвейер; производительность; сбоеустойчивость

DOI: 10.14357/08696527250108

EDN: QPTSAO

1 Введение

Самосинхронные схемы [1] образуют подкласс асинхронных схем [2], ставших альтернативой С-схемам [3] — современному базису реализации цифровых устройств. Асинхронные схемы реализуют дисциплину управления сигналами,

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

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

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

⁴Федеральный исследовательский центр «Информатика и управление» Российской академии наук, NMorozov@ipiran.ru

⁵Федеральный исследовательский центр «Информатика и управление» Российской академии наук, lplekhanov@inbox.ru

⁶Федеральный исследовательский центр «Информатика и управление» Российской академии наук, stepchenkov@mail.ru

операциями и процессами, не требующую использования глобальной синхронизации. Самосинхронные схемы в дополнение к этому контролируют успешное завершение всех переключений логических элементов, инициированных входными данными [1]. Тем самым они обеспечивают максимальную защиту от рас-синхронизации взаимодействия блоков в составе схемы, связанной с изменением задержек переключения логических элементов схемы из-за влияния окружающей среды и процессов старения.

Самосинхронные схемы характеризуются рядом преимуществ в сравнении с синхронными и асинхронными аналогами [4]:

- сохранением устойчивой работы при любых допустимых условиях эксплуатации вследствие независимости их функционирования от задержек переключения элементов их базиса реализации;
- гораздо более широким диапазоном работоспособности по напряжению питания и температуре окружающей среды;
- прекращением работы при появлении константной неисправности;
- естественной устойчивостью к логическим сбоям, вызываемым неблагоприятными внешними воздействиями;
- отсутствием аппаратных затрат и энергопотребления, связанных с реализацией глобального тактового «дерева».

Независимость от задержек переключения логических элементов и самодиагностика обеспечивают высокую эффективность создания надежных изделий, в том числе и отказоустойчивых, в базисе СС-схемотехники.

Основные недостатки СС-схем [5]:

- увеличенные аппаратные затраты из-за избыточного кодирования информации и схемы детектирования завершения всех переключений;
- наличие дополнительной задержки из-за двухфазной дисциплины функционирования и индикаторной подсхемы.

Аппаратная избыточность цифровой СС-схемы зависит от типа устройства. В сравнении с С-аналогами аппаратные затраты, выраженные в числе элементарных металл-оксид-полупроводник (КМОП) транзисторов, оказываются в 3,1 раза выше в комбинационных СС-схемах и в 1,5 раза больше в СС-схемах с памятью [6].

Однако последовательное уменьшение размеров транзисторов делает менее критичной избыточность аппаратных затрат. Дополнительное же время для переключения индикаторной подсхемы и пребывания СС-схемы в спейсерной фазе компенсируется отсутствием необходимости снижения производительности схемы в расчете на наихудший случай, поскольку СС-схемы работают в соответствии с реальными задержками элементов.

Интерес к СС-схемам повышается по мере нарастания проблем в С-схемотехнике, связанных с синхронизацией вычислительных процессов на большей площади топологической реализации микросхем. В печати приводятся результаты

разработки функционально-законченных асинхронных изделий уровня сверх-больших интегральных схем (СБИС) (см., например, [7]).

В СССР методология СС-проектирования активно развивалась коллективом исследователей под руководством В. И. Варшавского [1, 8]. С 1980-х гг. СС-методологию развивает ИПИ РАН [9–12].

Данная статья анализирует различные подходы к проектированию СС-схем, сравнивает характеристики С- и СС-аналогов, предлагает схемотехнические решения, повышающие быстродействие и сбоеустойчивость СС-схем, дает рекомендации по их использованию.

2 Практически значимые самосинхронные схемы

Анализ технических публикаций показывает, что большинство предлагаемых СС-решений на самом деле относятся к классу схем со связанными данными. Они используют запрос-ответное взаимодействие между функциональными блоками, но контроль окончания переключений заменен встроенными элементами задержки. Это уменьшает аппаратную избыточность и расширяет область работоспособности. Однако в результате теряются основные преимущества СС-схем — независимость их поведения от задержек элементов и самодиагностика.

Более практичным оказался подкласс схем, не зависящих от задержек в элементах (НЗЭ). В пределах эквихронной зоны [1] они обладают всеми свойствами действительно СС-схем. К устройствам такого типа относится, например, умножитель-сумматор (Fused Multiply-Add unit, FMA) [13].

Методология проектирования НЗЭ-схем, развиваемая в ИПИ РАН, базируется на двухфазном режиме работы, использовании парафазных, бифазных и унарных сигналов, контроле переключения всех элементов схемы индикаторной подсхемой, запрос-ответном взаимодействии СС-устройств.

Практически целесообразными оказались также квази-НЗЭ (КНЗЭ) схемы. Основное их отличие от НЗЭ-схем состоит в том, что они контролируют с помощью индикаторной подсхемы окончание переключения только тех элементов, которые стоят на критических путях обработки информации, а не всех элементов схемы. За счет этого они более быстродействующие и менее сложные, как, например, вычислитель [11], выполняющий деление и извлечение квадратного корня в соответствии со стандартом IEEE 754 [14]. Однако они обладают меньшим диапазоном работоспособности в сравнении с НЗЭ-схемами.

При микронных топологических нормах эквихронная зона измерялась миллиметрами и практически покрывала всю площадь кристалла БИС (большой интегральной схемы). Однако с переходом к субмикронным нормам размер эквихронной зоны существенно сократился из-за того, что задержки элементов стали меньше задержек в проводах. В современных цифровых СБИС эквихронная зона покрывает лишь малую часть площади кристалла, поэтому НЗЭ-схемы могут считаться СС-схемами, если функциональные блоки, связанные друг с другом

запрос-ответным взаимодействием, в топологии имеют соответствующие размеры.

К действительно СС-схемам относятся также схемы NULL Convention Logic (NCL) [15]. Они строятся на специальных элементах, каждый из которых индицирует все свои входы, и потому их разработка проще автоматизируется. Известны, например, две системы их синтеза: Balsa [16] и UNCLE [17]. Однако элементный базис NCL-схем существенно ограничен, а сами схемы чрезмерно избыточны и из-за этого обладают пониженным быстродействием и высоким энергопотреблением. NCL-схемы оказываются сложнее НЗЭ-схем в 4,5 раза при реализации, например, двоичного счетчика, в 1,13 раза при реализации умножителя и не более чем в 2 раза при реализации простых логических схем [18].

3 Программная поддержка

Маршрут проектирования НЗЭ-схем поддерживается отечественными программами анализа разрабатываемой схемы на самосинхронность [19, 20]. Они обеспечивают проверку схемы на принадлежность к классу СС-схем и помогают пользователю найти ошибки в схеме.

В настоящее время разрабатывается подсистема синтеза НЗЭ-схем [21] на основе библиотек стандартных элементов и СС-ячеек [22, 23]. Ряд триггеров СС-библиотеки обеспечивают эффективное решение проблем:

- увеличения нагрузочной способности выходов триггеров [24];
- сокращения числа сигналов входного интерфейса [25];
- реализации триггера с парафазными входом и выходом [26].

Модели библиотечных элементов в системе автоматизированного проектирования (САПР) «Ковчег» генерируют предупреждения о возможном нарушении дисциплины формирования сигналов. Это позволяет разработчику НЗЭ-схемы на этапе моделирования избежать ошибок при разработке схемы.

4 Экспериментальные результаты

Сравнение представленных ниже С-, КНЗЭ- и НЗЭ-вариантов реализации цифровых устройств основано на оценке быстродействия в реальных условиях эксплуатации для СС-схем и в наихудшем случае для С-схем. Результаты испытаний ряда цифровых устройств, описанные ниже, подтверждают, что в нормальных условиях СС-схемы, как правило, быстрее С-аналогов.

4.1 Микроядро

Микроядро представляет собой 4-разрядный аналог вычислительного ядра микроконтроллера PIC18XX (Microchip). Его структурная схема показана на

рис. 1 [27]. Схема микроядра была разработана в С- и НЗЭ-вариантах и изготовлена на базовом матричном кристалле (БМК) серии 5503 по 1,6-микрометровой КМОП-технологии (НПК «Технологический центр», МИЭТ). Результаты измерений длительности выполнения разных наборов из четырех команд для обоих вариантов микроядра представлены в табл. 1. Они показывают, что НЗЭ-вариант микроядра лучше своего С-аналога в 3,25 раза по реальной производительности. Таблица 2 описывает сложность аппаратных блоков микроядра для обоих вариантов в вентильях БМК. Для С-варианта они оказались меньше в 1,43 раза.

Рисунок 2 показывает результаты проверки работоспособности С- (тем-

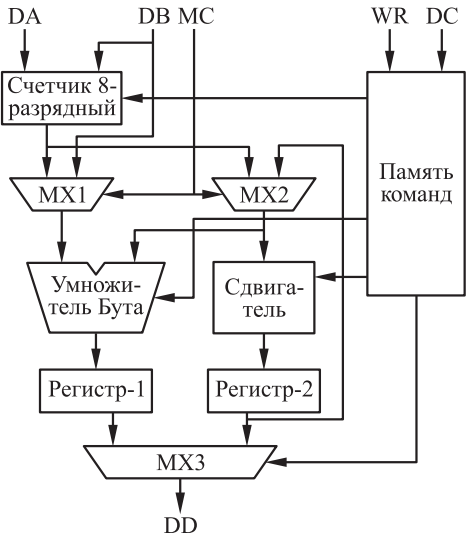


Рис. 1 Структурная схема микроядра

Таблица 1 Производительность микроядра

	Набор операций	Производительность, нс			
		С-вариант	НЗЭ-вариант		
		$U = 4,5 \text{ В}, T = 125 \text{ }^{\circ}\text{C}$	$U = 4,5 \text{ В}, T = 125 \text{ }^{\circ}\text{C}$	$U = 5,0 \text{ В}, T = 27 \text{ }^{\circ}\text{C}$	$U = 5,5 \text{ В}, T = -63 \text{ }^{\circ}\text{C}$
1	Cyclic MUL	250	166	144	118
2	Cyclic ROT	250	121	102	86
3	Cyclic NOP	250	111	93	75
4	Cyclic JUMP	500	90	78	66
5	MUL + JUMP + NOP + ROT	1248	516	440	364

Таблица 2 Сложность микроядра

	Аппаратный блок	С-вариант	НЗЭ-вариант
1	Умножитель	177	444
2	Сдвигатель	52	214
3	Счетчик	88	159
4	Память команд	230	192
5	Схема управления	423	380
Итого		970	1389

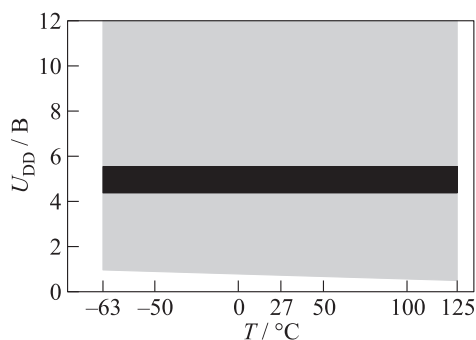


Рис. 2 Диапазоны работоспособности С- и НЗЭ-варианта микродра

ная область) и НЗЭ-образцов микродра в диапазоне напряжения питания U_{DD} и температуры T . Они подтверждают, что диапазон работоспособности НЗЭ-образцов микродра в несколько раз шире гарантированного диапазона образцов С-аналога.

4.2 Сопроцессор

Сопроцессор [11] выполняет операции деления и извлечения квадратного корня над числами одинарной и двойной точности. Схема сопроцессора, реализованного в виде вариантов КНЗЭ и НЗЭ, показана на рис. 3. Она содержит следующие блоки: входной регистр операндов и признаков операции (ВР); блок

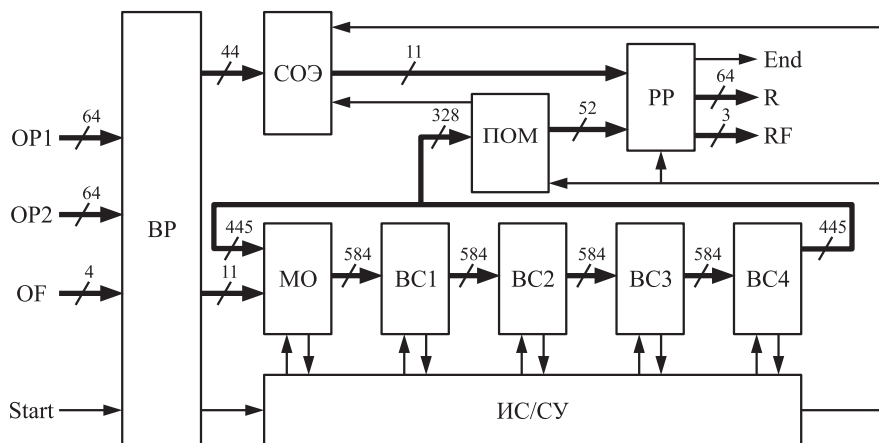


Рис. 3 Структурная схема сопроцессора

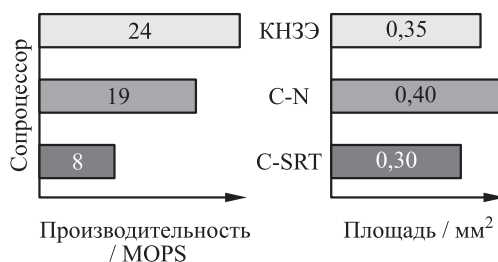


Рис. 4 Характеристики вариантов сопроцессоров

обработки экспонент (СОЭ); входной мультиплексор операндов (МО), организующий итерационное выполнение операции; четыре однотипные вычислительные стадии ВС1–ВС4, каждая из которых вычисляет один бит результата операции; индикаторную схему (ИС) со схемой управления (СУ), обеспечивающую СС-дисциплину обмена данными между блоками вычислителя; блок постобработки мантиссы (ПОМ); выходной регистр результата (РР).

Для сравнения использовались четыре варианта сопроцессора: синхронный C-N-вариант, реализующий алгоритм Ньютона; синхронный C-SRT-вариант, реализующий алгоритм SRT Radix4; КНЗЭ- и НЗЭ-варианты, реализующие алгоритм SRT Radix2. Первые три варианта были реализованы в составе микросхемы по стандартной 0,18-микрометровой КМОП-технологии. Анализ параметров трех вариантов показывает, что КНЗЭ-реализация обладает лучшей производительностью и площадью топологической реализации, как показано на рис. 4.

В сравнении с НЗЭ-вариантом КНЗЭ-вариант сопроцессора имеет лучшее на 30%–40% быстродействие и меньшую на 20% сложность. В табл. 3 приведено усредненное время выполнения деления (DIV) и извлечения квадратного корня (SQRT) с двойной точностью КНЗЭ- и НЗЭ-сoproцессором по результатам моделирования статистически значимого набора операндов.

Анализ табл. 3 показывает, что оба варианта сопроцессора продолжают устойчиво работать при понижении напряжения питания до 0,32 В при номинальном

Таблица 3 Производительность КНЗЭ и НЗЭ вариантов сопроцессора

	Условия		КНЗЭ		НЗЭ	
	U_{DD} , В	T , °С	DIV, нс	SQRT, нс	DIV, нс	SQRT, нс
1	1,98	–63	34,7	36,9	47,3	50,2
2	1,80	25	46,7	49,1	63,5	67,0
3	1,62	125	63,9	70,3	86,9	90,1
4	0,32	125	25 688	25 301	34 940	34 410
5	0,20	125	—	—	340 800	336 920

уровне 1,8 В. При этом быстродействие падает в 550 раз, но все операции выполняются корректно. Это свидетельствует о широкой зоне работоспособности КНЗЭ- и НЗЭ-сопроцессора. Однако только НЗЭ-сопроцессор оказывается работоспособным при более низком напряжении питания в 0,2 В вследствие использования в нем полной индикации, соответствующей СС-схемам.

5 Ускорение индикации самосинхронных схем

Основными факторами ограничения быстродействия СС-схем считаются двухфазная дисциплина работы и наличие индикаторной подсхемы. Спейсерная фаза разделяет два соседних рабочих состояния схемы, а для подтверждения окончания переключения СС-схемы в каждую фазу нужно проиндицировать все элементы схемы и сформировать один индикаторный сигнал.

5.1 Задержка индикаторной подсхемы

В схемах с небольшой разрядностью обрабатываемых данных индикаторная подсхема работает в фоновом режиме и слабо влияет на быстродействие схемы. В многоразрядных СС-схемах большое число индицирующих сигналов приводит к увеличению числа каскадов индикаторной подсхемы, и она начинает вносить существенный вклад в задержку СС-схемы. В ряде случаев ее задержка может быть уменьшена за счет распараллеливания индикации и управления.

Сложные СС-схемы, как и их С-аналоги, строятся в виде конвейера. Базовый принцип его работы состоит в следующем: k -я ступень конвейера начинает переключаться в рабочую (спейсерную) фазу только после того, как $(k - 1)$ -я ступень переключилась в рабочую (спейсерную) фазу, а $(k + 1)$ -я ступень переключилась в спейсерную (рабочую) фазу. На рис. 5 приведена схема типового СС-конвейера. Здесь «Логика» — комбинационная часть ступени конвейера; «Регистр» — выходной регистр ступени; «С» — С-элемент Маллера [2] (гистерезисный триггер [1]), формирующий сигнал управления регистром. Входами С-элемента служат индикаторные выходы комбинационной части и регистра ступени конвейера.

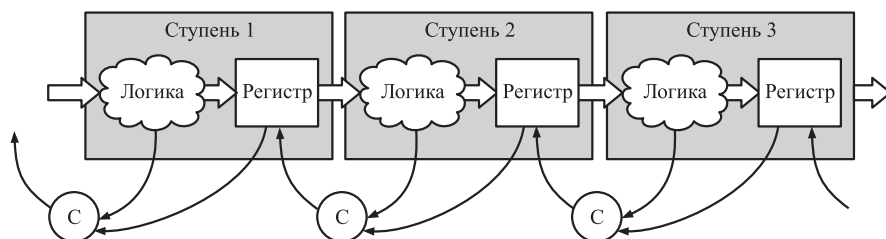


Рис. 5 Структурная схема типового СС-конвейера

Сборка частичных индикаторов в один общий сигнал реализуется пирамидой С-элементов. Задержка сборки общих индикаторных выходов «Логике» и «Регистра» ступени равна

$$t_{IC} = t_{C3} \cdot \lceil \log_3 M \rceil + t_{C2},$$

где t_{C2} — задержка двухвходового С-элемента; t_{C3} — задержка трехвходового С-элемента; M — число разрядов или выходов «Логике» и «Регистра».

Фактически t_{IC} становится дополнительной задержкой переключения ступени конвейера. Логарифмическая зависимость задержки от числа разрядов схемы смягчает влияние индикаторной подсхемы на быстродействие СС-схемы, но в многоразрядных вычислительных устройствах она играет драматическую роль. Например, число индикаторных сигналов первого каскада парафазного «дерева» Уоллеса 64-разрядного умножителя [28] равно 1430. Индикаторная подсхема реализована на 716 двух- и трехвходовых С-элементах в виде 7-каскадной пирамидальной структуры. В 65-нанометровой КМОП-технологии типовая задержка одного трехвходового С-элемента равна 50 пс. В результате задержка переключения такой индикаторной схемы составляет около 350 пс, что приводит к серьезному замедлению конвейера умножителя.

Однако для СС-схем с высокой параллельностью обработки данных существует решение проблемы задержки индикаторной подсхемы.

5.2 Оптимизация индикаторной подсхемы

Использование поразрядной или групповой индикации и управления обеспечивает сокращение вклада индикаторной схемы в задержку переключения многоразрядной СС-схемы [29]. Рисунок 6 иллюстрирует структурную схему СС-конвейера с поразрядной индикацией.

Поразрядные индикаторные сигналы «Логике» интегрируются в «Регистр» для формирования поразрядных индикаторов ступени конвейера. Рисунок 7 показывает два варианта интеграции индикатора «Логике» I_L в индикатор I_R разряда «Регистра». Здесь X , XB — парафазный информационный выход

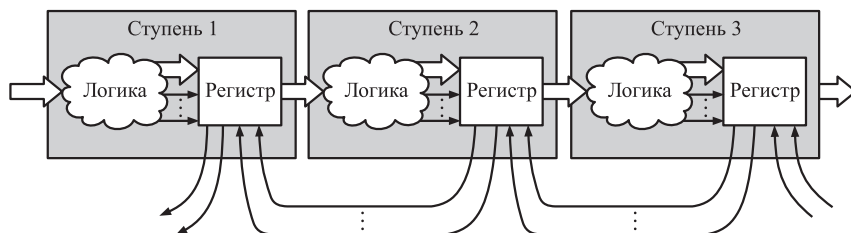


Рис. 6 Структурная схема СС-конвейера с поразрядной индикацией

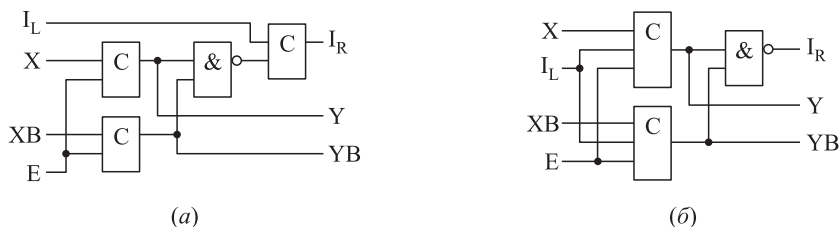


Рис. 7 Схема разряда «Регистра» на двухвходовых (а) и трехвходовых (б) С-элементах

«Логике»; Е — общий или поразрядный вход управления регистра; Y, YB — парафазный информационный выход «Регистра».

Поразрядная индикация позволяет инициировать переключение отдельных входов ступени СС-конвейера в следующую фазу работы, если все выходы «Регистра», зависящие от этих входов, уже переключились в текущую фазу.

Пусть СС-схема имеет M разрядов и каждые K соседних разрядов связаны друг с другом сигналами. Назовем K коэффициентом связности. Тогда индикация окончания переключения K соседних разрядов в очередную фазу работы гарантирует готовность самого младшего из них к переключению в противоположную фазу. Для полной индикации всех M разрядов ступени конвейера потребуется $(M - K + 1)$ таких индикаторных подсхем.

Максимальный эффект от использования поразрядной индикации достигается в СС-схемах с абсолютной независимостью соседних разрядов ($K = 1$). При увеличении K растут аппаратные затраты и снижается быстродействие из-за формирователей групповых индикаторов.

Аппаратные затраты поразрядной индикации избыточны, поскольку один и тот же разряд схемы индицируется одновременно в K индикаторных подсхемах, и равны примерно $C_K = 16(M - K + 1) \log_3 K$ КМОП-транзисторов. Если выполняется соотношение $K < L = (\lceil \log_3 K \rceil)^3$, то для минимизации аппаратных затрат надо индицировать одним общим сигналом L соседних разрядов. Тогда аппаратные затраты снижаются до уровня

$$C_L = 16 \frac{M - L + 1}{L - K + 1} \log_3 L.$$

Например, при $M = 64$, $K = 5$ и $L = 9$ сложность индикаторной схемы снижается в 3,9 раза. В этом случае индикатор становится групповым, а не поразрядным.

В качестве примера рассмотрим «дерево» Уоллеса в умножителе Бута. Оно состоит из многоразрядных сумматоров с сохранением переноса. Коэффициент связности дерева Уоллеса равен числу его каскадов. Для парафазного дерева Уоллеса [28] $K = 7$. Разбиение дерева Уоллеса на две ступени конвейера с тремя

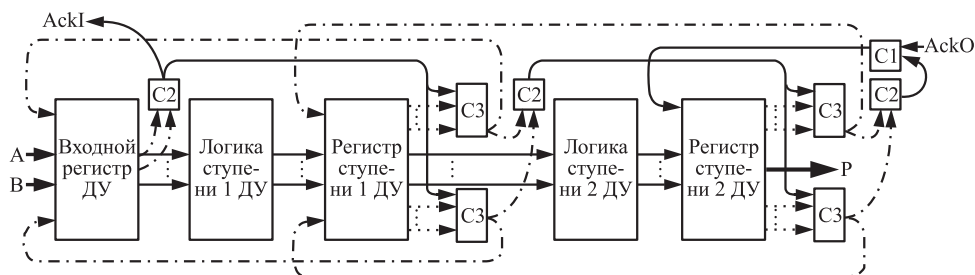


Рис. 8 Структурная схема НЗЭ-конвейера с групповой индикацией

Таблица 4 Параметры вариантов индикации «дерева» Уоллеса

	Вариант индикации	Длительность цикла, пс	Число КМОП-транзисторов
1	Классический	970	220 000
2	Групповой	710	225 500

и четырьмя каскадами соответственно уменьшает коэффициент связности для каждой ступени до величин $K_1 = 3$ и $K_2 = 4$.

Рисунок 8 иллюстрирует структурную схему двухступенчатого конвейера дерева Уоллеса. Здесь A и B — сомножители; P — произведение; AckI и AckO — входной и выходной сигналы уведомления об окончании текущей фазы работы. Индикаторная подсхема «C1» формирует сигнал управления фазой выходного регистра. Подсхемы «C2» объединяют групповые индикаторы в один общий индикатор. Подсхемы «C3» формируют групповые индикаторные сигналы.

Число входов в «C3» равно коэффициенту зависимости ступени плюс еще один, куда подается общий индикатор регистра, входами которого управляют выходы «C3», для обеспечения действительно СС-работы.

В табл. 4 приведены результаты моделирования парафазного дерева Уоллеса с классической и групповой индикацией для 65-нанометровой КМОП-технологии. Средняя длительность рабочего цикла варианта с групповым управлением оказывается на 37% меньше, чем в классическом варианте. Платой за это служит увеличение аппаратных затрат на 2,5%.

Использование групповой индикации ограничено только внутренней областью СС-конвейера. Входная и выходная ступени конвейера обязаны формировать общий индикаторный сигнал для всех разрядов регистра.

6 Повышение быстродействия самосинхронного конвейера

Как известно, рост производительности С-схемы достигается за счет ее конвейеризации, увеличения числа ступеней конвейера и внедрения параллельных и спекулятивных вычислений. Это справедливо и для СС-схем.

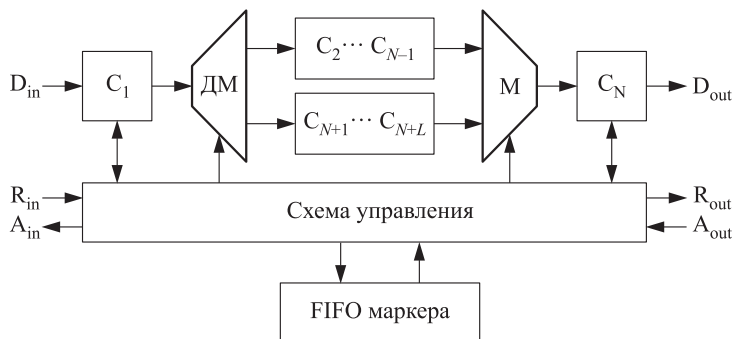


Рис. 9 Структурная схема мультиплексируемого СС-конвейера

На рис. 9 представлен пример СС-конвейера с мультиплексируемыми путями [30]. Он содержит N ступеней в основном тракте обработки данных и L ($L < N - 2$) ступеней в параллельном тракте. Здесь D_{in} — входные данные; D_{out} — выходные данные; C_k , $k = \overline{1, N+L}$, — ступени конвейера; ДМ — демультиплексор данных 1 : 2; М — мультиплексор данных 2 : 1; R_{in} и R_{out} — входной и выходной сигналы запроса; A_{in} и A_{out} — входной и выходной сигналы подтверждения. При выполнении некоторого условия данные направляются в более короткий тракт, и вычислительный процесс ускоряется.

Блок FIFO (First-Input-First-Output) маркера обеспечивает соблюдение порядка обработки данных ступенью CN. Самосинхронный конвейер функционирует корректно при любой емкости FIFO.

Однако для оптимального функционирования конвейера на рис. 9 FIFO должно иметь емкость больше $0,5(N + L)$ бит.

Предложенный способ мультиплексирования тракта обработки данных обеспечивает оптимальную реализацию СС-конвейера. Он позволяет повысить среднюю производительность конвейера, если задержка одного из его параллельных трактов существенно меньше задержки другого тракта.

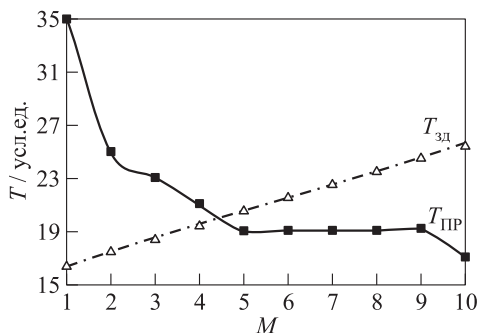


Рис. 10 Зависимость производительности СС-конвейера от числа ступеней

Производительность СС-конвейера также зависит от числа его ступеней. Следует, однако, учесть, что разбиение СС-схемы на ступени конвейера сопровождается добавлением регистров и их индикаторных подсхем, вносящих дополнительную задержку. Рисунок 10 [31] демонстрирует зависимости периода появления результата на выходе СС-конвейера $T_{пр}$ и времени прохода

данных $T_{зд}$ по нему от числа ступеней M , на которое разбивается исходная схема. Оптимальное с точки зрения быстродействия число ступеней СС-конвейера зависит от ожидаемого темпа поступления данных на вход конвейера.

В отличие от С-конвейера, производительность СС-конвейера определяется двумя самыми медленными ступенями: четной и нечетной.

7 Повышение сбоеустойчивости самосинхронных схем

Благодаря избыточному кодированию информации и двухфазной дисциплине работы [1], СС-схемы более устойчивы к воздействию дестабилизирующих факторов и помех, чем их С-аналоги [32]. Сравнительный анализ на эмпирическом уровне показал, что время бессбойной работы СС-схем превышает аналогичный показатель С-схем в 4 раза в комбинационных схемах и в 7,1 раза в последовательностных [6]. Математические модели появления критического сбоя в С- и СС-конвейере, объединяющем в себе комбинационные и последовательностные схемы, подтверждают этот вывод [33].

Способы повышения сбоеустойчивости СС-конвейера включают в себя:

- использование на первом каскаде индикаторной подсхемы элементов равнозначности, детектирующих антиспейсерное состояние парафазного сигнала как спейсерное [32];
- реализацию разряда регистра ступени на С-элементе с защитой записи антиспейсера и от залипания в антиспейсере [34];
- разрешение записи в регистр конвейера только после подтверждения завершения переключения регистра предыдущей ступени [34];
- использование DICE-подобных С-элементов [35] в индикаторной подсхеме.

8 Автоматизация проектирования самосинхронных схем

Внедрение СС-схем в практические разработки современных СБИС невозможно без использования программных средств, обеспечивающих логический синтез и проверку схемы на самосинхронность. Известные системы синтеза СС-схем (например, Balsa [16] и UNCLE [17]) требуют ввода исходного описания синтезируемой схемы в специальных форматах, предполагающих наличие у разработчика достаточно высокой квалификации в области СС-схем. К тому же они не всегда гарантируют получение адекватного результата [2].

В ФИЦ ИУ РАН имеются средства САПР, обеспечивающие СС-анализ схемы [19, 20], синтез комбинационных СС-схем [36] и их характеризацию [37]. Они обеспечивают корректную разработку СС-схем разной сложности и формирование библиотеки СС-элементов для интеграции в промышленные САПР СБИС.

В настоящее время в ФИЦИУ РАН создаются программные средства синтеза последовательностных и конвейерных СС-схем и разрабатывается методология и средства синтеза СС-схем любого уровня сложности на основе исходного синхронного описания на языке Verilog [38]. В рамках этой работы будет решен вопрос оптимального синтеза СС-схем с памятью.

9 Выводы

Самосинхронные схемы в сравнении с С-аналогами имеют целый ряд преимуществ, обеспечивающих их работоспособность в гораздо более широком диапазоне условий эксплуатации. Однако они характеризуются большей сложностью комбинационных схем (в 3,1 раза) и последовательностных реализаций (в 1,5 раза). Применение СС-схемотехники оправдано в областях, где высокая стабильность работы выступает главным критерием.

Типовые вычислительные СС-устройства малой разрядности в 1,7–2,6 раза лучше С-аналогов по производительности в реальных условиях, что подтверждается результатами измерений СС-схем разных типов и сложности.

Применение поразрядной или групповой индикации и управления в многоразрядных схемах позволяет существенно ускорить их работу за счет относительно небольшого увеличения аппаратных затрат в том случае, если соседние разряды схемы слабо связаны друг с другом общими сигналами. Например, быстродействие СС-конвейера дерева Уоллеса парафазного умножителя Бута повышается на 37% при увеличении сложности на 2,5%.

Самосинхронные схемы обладают в несколько раз более высоким уровнем устойчивости к однократным логическим сбоям, чем их С-аналоги. Предложенные схемотехнические методы дополнительно повышают сбоеустойчивость комбинационных и конвейерных СС-схем.

Разрабатываемые средства автоматизации проектирования СС-схем создают основу для широкого внедрения СС-схем в практические изделия микроэлектронной техники, ставящие на первое место надежность работы.

Литература

1. Varshavsky V., Kishinevsky M., Marakhovsky V., Peschansky V.A. Self-timed control of concurrent processes / Ed. V. Varshavsky. — Kluwer Academic Publs., 1990. 245 p.
2. Sparsø J. Introduction to asynchronous circuit design. — Copenhagen, Denmark: DTU Compute, Technical University of Denmark, 2020. 269 p. https://backend.orbit.dtu.dk/ws/portalfiles/portal/215895041/JSPA_async_book_2020_PDF.pdf.
3. Harris D., Harris S. L. Digital design and computer architecture. — Elsevier, 2013. 690 p.
4. Zakharov V., Stepchenkov Yu., Diachenko Yu., Rogdestvenski Yu. Self-timed circuitry retrospective // Conference (International) on Engineering Technologies and

- Computer Science Proceedings. — Piscataway, NJ, USA: IEEE, 2020. P. 58–64. doi: 10.1109/EnT48576.2020.00018.
5. Плеханов Л. П. Основы самосинхронных электронных схем. — М.: Бином; Лаборатория знаний, 2013. 208 с. EDN: SUMKIV.
 6. Соколов И. А., Степченко Ю. А., Рождественский Ю. В., Дьяченко Ю. Г. Приближенная оценка эффективности синхронной и самосинхронной методологий в задачах проектирования сбоеустойчивых вычислительно-управляющих систем // Автоматика и телемеханика, 2022. № 2. С. 122–132. doi: 10.31857/S0005231022020088. EDN: PYRRXN.
 7. Bink A., York R. ARM996HS: The first licensable, clockless 32-bit processor core // IEEE Micro, 2007. Vol. 27. No. 2. P. 58–68. doi: 10.1109/MM.2007.28.
 8. Kishinevsky M., Kondratyev A., Taubin A., Varshavsky V. Concurrent hardware: The theory and practice of self-timed design. — New York, NY, USA: J. Wiley & Sons, 1994. 368 p.
 9. Плеханов Л. П., Степченко Ю. А. Экспериментальная проверка некоторых свойств строго самосинхронных схем // Системы и средства информатики, 2006. Вып. 16. С. 476–485. EDN: KZUWOX.
 10. Соколов И. А., Степченко Ю. А., Петрухин В. С., Дьяченко Ю. Г., Захаров В. Н. Самосинхронная схемотехника — перспективный путь реализации аппаратуры // Системы высокой доступности, 2007. Т. 3. № 1-2. С. 61–72.
 11. Stepchenkov Yu., Diachenko Yu., Zakharov V., Rogdestvenski Yu., Morozov N., Stepchenkov D. Quasi-delay-insensitive computing device: Methodological aspects and practical implementation // Integrated circuit and system design: Power and timing modeling, optimization and simulation / Eds. J. Monteiro, R. Leuken. — Lecture notes in computer science ser. — Berlin, Heidelberg: Springer, 2010. Vol. 5953. P. 276–285. doi: 10.1007/978-3-642-11802-9_32
 12. Степченко Ю. А., Дьяченко Ю. Г., Рождественский Ю. В., Морозов Н. В., Степченко Д. Ю. Разработка вычислителя, не зависящего от задержек элементов // Системы и средства информатики, 2010. Вып. 20. № 1. С. 5–23. EDN: NTMGCH.
 13. Stepchenkov Yu., Rogdestvenski Yu., Diachenko Yu., Stepchenkov D., Shikunov Yu. Energy efficient speed-independent 64-bit fused multiply-add unit // Conference of Russian Young Researchers in Electrical and Electronic Engineering Proceedings. — IEEE, 2019. P. 1709–1714. doi: 10.1109/EIConRus.2019.8657207.
 14. IEEE Computer Society. IEEE standard for floating-point arithmetic 754-2008, 2008. 70 p. doi: 10.1109/IEEESTD.2008.4610935.
 15. Fant K. M. Logically determined design: Clockless system design with NULL convention logic. — New York, NY, USA: John Wiley, 2005. 292 p. doi: 10.1002/0471702897.
 16. Edwards D., Bardsley A., Jani L., Plana L., Toms W. Balsa: A tutorial guide. Version V3.5. — Manchester, 2006. 157 p. <https://apt.cs.manchester.ac.uk/ftp/pub/amulet/balsa/3.5/BalsaManual3.5.pdf>.
 17. Tailor R. A., Reese R. B. UNCLE — Unified NCL Environment — an NCL design tool // Asynchronous circuit applications. — Institute of Engineering and Technology, 2019. P. 293–307. doi: 10.1049/PBCS061E_ch14.
 18. Соколов И. А., Степченко Ю. А., Бобков С. Г., Захаров В. Н., Дьяченко Ю. Г., Рождественский Ю. В., Сурков А. В. Базис реализации супер-ЭВМ эксафлопсного

- класса // Информатика и её применения, 2014. Т. 8. Вып. 1. С. 45–70. doi: 10.14357/19922264140106. EDN: RYYFGD.
19. Рождественский Ю. В., Морозов Н. В., Степченков Ю. А., Рождественские А. В. Универсальная подсистема анализа самосинхронных схем // Системы и средства информатики, 2006. Т. 16. № 1. С. 463–475. EDN: KZUWON.
 20. Рождественский Ю. В., Морозов Н. В., Рождественские А. В. Подсистема событийного анализа самосинхронных схем АСПЕКТ // Проблемы разработки перспективных микро- и наноэлектронных систем. — М.: ИППМ РАН, 2010. С. 26–31.
 21. Плеханов Л. П., Денисов А. Н., Дьяченко Ю. Г., Степченков Ю. А., Мамонов Д. И., Степченков Д. Ю. Синтез самосинхронных схем в базисе БМК // 5-я Междунар. научн. конф. «Электронная компонентная база и микроэлектронные модули»: Сб. тезисов. — М.: Техносфера, 2019. С. 450–454.
 22. Stepchenkov Yu. A., Zakharov V. N., Diachenko Yu. G., Morozov N. V., Stepchenkov D. Yu. Cell library for speed-independent VLSI // IEEE East-West Design and Test Symposium Proceedings. — IEEE, 2015. P. 137–140. doi: 10.1109/EWDTS.2015.7493111.
 23. Степченков Ю. А., Денисов А. Н., Дьяченко Ю. Г. и др. Библиотека функциональных ячеек для проектирования самосинхронных полужаказных БМК микросхем серий 5503/5507. — М.: Техносфера, 2017. Т. 4. 376 с.
 24. Sokolov I. A., Stepchenkov Y. A., Dyachenko Y. G. Self-timed RS-trigger with the enhanced noise immunity. U.S. Patent No. 8232825, 2012. 31 p.
 25. Sokolov I. A., Stepchenkov Y. A., Dyachenko Y. G. Self-timed trigger with single-rail data input. U.S. Patent No. 8324938, 2012. 34 p.
 26. Захаров В. Н., Степченков Ю. А., Дьяченко Ю. Г., Плеханов Л. П., Морозов Н. В., Дьяченко Д. Ю. Самосинхронный двухтактный триггер с парафазными входными и выходными сигналами с нулевым спейсером. Патент на изобретение № 2835382 с приоритетом от 13.09.24. Оpubл. 25.02.2025, бюл. № 6.
 27. Степченков Ю. А., Петрухин В. С., Дьяченко Ю. Г. Опыт разработки самосинхронного ядра микроконтроллера на базовом матричном кристалле // Нано- и микросистемная техника, 2006. № 5. С. 29–36. EDN: IAGLLN.
 28. Stepchenkov Yu. A., Diachenko Yu. G., Rogdestvenski Yu. V., Diachenko D. Yu., Shikunov Yu. I. Self-timed multiply-add-subtract unit alternates // Conference of Russian Young Researchers in Electrical and Electronic Engineering Proceedings. — IEEE, 2020. P. 1864–1868. doi: 10.1109/EIConRus49466.2020.9039039.
 29. Степченков Ю. А., Дьяченко Ю. Г., Рождественский Ю. В. и др. Оптимизация индикации многоразрядных самосинхронных схем // Системы и средства информатики, 2019. Т. 29. № 4. С. 14–27. doi: 10.14357/08696527190402. EDN: XFXKVZ.
 30. Степченков Ю. А., Дьяченко Ю. Г., Степченков Д. Ю. и др. Мультиплексируемый самосинхронный конвейер // Системы и средства информатики, 2023. Т. 33. № 2. С. 4–12. doi: 10.14357/08696527230201. EDN: FLVEWP.
 31. Степченков Ю. А., Дьяченко Ю. Г., Морозов Н. В. и др. Оптимизация самосинхронного конвейера // Системы высокой доступности, 2023. Т. 19. № 1. С. 5–13. doi: 10.18127/j20729472-202301-01. EDN: IRTFWY.

32. *Stepchenkov Yu. A., Kamenskih A. N., Diachenko Yu. G., Rogdestvenski Yu. V., Diachenko D. Yu.* Improvement of the natural self-timed circuit tolerance to short-term soft errors // *Advances Science Technology Engineering Systems J.*, 2020. Vol. 5. Iss. 2. P. 44–56. doi: 10.25046/aj050206.
33. *Sokolov I., Stepchenkov Yu., Diachenko Yu., Khilko D.* Mathematical models of critical soft error in synchronous and self-timed pipeline // *Mathematics*, 2025. Vol. 13. Iss. 5. Art. 695. 15 p. doi: 10.3390/math13050695.
34. *Stepchenkov Yu. A., Diachenko Yu. G., Appolonov G. S., et al.* Self-timed pipeline protection against critical fault // *Conference of Young Researchers in Electrical and Electronic Engineering Proceedings*. — IEEE, 2025. P. 179–182.
35. *Соколов И. А., Степченков Ю. А., Дьяченко Ю. Г. и др.* Анализ сбоеустойчивости самосинхронного конвейера // *Системы и средства информатики*, 2022. Т. 32. № 4. С. 4–13. doi: 10.14357/08696527220401. EDN: NXADJN.
36. *Плеханов Л. П.* Программа синтеза комбинационных схем на заданной библиотеке элементов СИНТАБИБ. Свидетельство о госрегистрации программы для ЭВМ № 2007613665 от 22.08.2007.
37. *Морозов Н. В., Дьяченко Ю. Г., Степченков Д. Ю., Рождественские А. В.* Система характеристики самосинхронных элементов САХИБ. Версия 4. Свидетельство о госрегистрации программы для ЭВМ № 2021668787 от 19.11.2021.
38. *Зацаринный А. А., Степченков Ю. А., Дьяченко Ю. Г. и др.* Автоматизация синтеза самосинхронных схем // *Системы высокой доступности*, 2023. Т. 19. № 3. С. 48–56. doi: 10.18127/j20729472-202303-04. EDN: IAPVXR.

Поступила в редакцию 27.12.2024

Принята к публикации 15.02.2025

PROPERTIES AND OPTIMIZATION OF SELF-TIMED CIRCUITS

*V. N. Zakharov, Yu. A. Stepchenkov, Yu. G. Diachenko, N. V. Morozov,
L. P. Plekhanov, and D. Yu. Stepchenkov*

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

Abstract: The article is devoted to the brief overview of the development trends and achievements of self-timed (ST) circuits. Due to redundant data coding and two-phase discipline, ST circuits have a number of advantages over synchronous counterparts. They operate reliably with any cell delays, detect any constant faults, and have an extremely wide performance range in terms of supply voltage and ambient temperature. Practical development of ST units with varying complexity has proven the effectiveness of ST solutions, especially in highly reliable and fault-tolerant applications. The article presents the comparative test results of synchronous and ST circuit test chips and the performance and fault tolerance estimates for ST circuits of varying complexity. Group indication of multibit ST circuits increases their performance by 40% due to a slight

increase (less than 3%) in hardware complexity. The developed design methods of protection against soft errors improve the natural immunity of ST circuits to adverse effects and ensure a level of fault tolerance of ST circuits several times higher than that of the synchronous counterparts.

Keywords: self-timed circuit; C-element; dual-rail; indication; pipeline; performance; fault tolerance

DOI: 10.14357/08696527250108

EDN: QPTSAO

References

1. Varshavsky, V. I., M. A. Kishinevsky, V. B. Marakhovsky, and V. A. Peschansky. 1990. *Self-timed control of concurrent processes*. Kluwer Academic Publs. 245 p.
2. Sparsø, J. 2020. *Introduction to asynchronous circuit design*. Copenhagen, Denmark: DTU Compute, Technical University of Denmark. 269 p. Available at: https://backend.orbit.dtu.dk/ws/portalfiles/portal/215895041/JSPA_async_book_2020.PDF.pdf (accessed March 6, 2025).
3. Harris, D., and S. L. Harris. 2013. *Digital design and computer architecture*. Elsevier. 690 p.
4. Zakharov, V., Yu. Stepchenkov, Yu. Diachenko, and Yu. Rogdestvenski. 2020. Self-timed circuitry retrospective. *Conference (International) on Engineering Technologies and Computer Science Proceedings*. Piscataway, NJ: IEEE. 58–64. doi: 10.1109/EnT48576.2020.00018.
5. Plekhanov, L. P. 2013. *Osnovy samosinkhronnykh elektronnykh skhem* [Basics of self-timed electronic circuits]. Moscow: BINOM; Laboratoriya znaniy. 208 p. EDN: SUMKIV.
6. Sokolov, I. A., Yu. A. Stepchenkov, Yu. V. Rogdestvenski, and Yu. G. Diachenko. 2022. Approximate evaluation of the efficiency of synchronous and self-timed methodologies in problems of designing failure-tolerant computing and control systems. *Automat. Rem. Contr.* 83(2):264–272. doi: 10.1134/S0005117922020084. EDN: CDSNSD.
7. Bink, A., and R. York. 2007. ARM996HS: The first licensable, clockless 32-bit processor core. *IEEE Micro* 27(2):58–68. doi: 10.1109/MM.2007.28.
8. Kishinevsky, M., A. Kondratyev, A. Taubin, and V. Varshavsky. 1994. *Concurrent hardware: The theory and practice of self-timed design*. New York, NY: John Wiley & Sons. 368 p.
9. Plekhanov, L. P., and Yu. A. Stepchenkov. 2006. Eksperimental'naya proverka nekotorykh svoystv strogo samosinkhronnykh skhem [Experimental verification of some properties of strictly self-timed circuits]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 16:476–485. EDN: KZUWOX.
10. Sokolov, I. A., Yu. A. Stepchenkov, V. S. Petrukhin, Yu. G. Diachenko, and V. N. Zakharov. 2007. Samosinkhronnaya skhemotekhnika — perspektivnyy put' realizatsii apparatury [Self-timed circuitry is the perspective way for hardware realization]. *Sistemy vysokoy dostupnosti* [Highly Available Systems] 3(1-2):61–72.
11. Stepchenkov, Yu., Yu. Diachenko, V. Zakharov, Yu. Rogdestvenski, N. Morozov, and D. Stepchenkov. 2009. Quasi-delay-insensitive computing device: Methodological aspects and practical implementation. *Integrated circuit and system design: Power*

- and timing modeling, optimization and simulation*. Eds. J. Monteiro and R. Leuken. Lecture notes in computer science ser. Berlin, Heidelberg: Springer. 5953:276–285. doi: 10.1007/978-3-642-11802-9_32.
12. Stepchenkov, Yu. A., Yu. G. Diachenko, Yu. V. Rogdestvenski, N. V. Morozov, and D. Yu. Stepchenkov. 2010. Razrabotka vychislitelya, nezavisyashchego ot zaderzhek elementov [Designing of the delay-independent computing device]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 20(1):5–23. EDN: NTMGCH.
13. Stepchenkov, Yu. A., Yu. V. Rogdestvenski, Yu. G. Diachenko, D. Yu. Stepchenkov, and Yu. Shikunov. 2019. Energy efficient speed-independent 64-bit fused multiply-add unit. *Conference of Russian Young Researchers in Electrical and Electronic Engineering Proceedings*. IEEE. 1709–1714. doi: 10.1109/EIConRus.2019.8657207.
14. IEEE Computer Society. 2008. IEEE standard for floating-point arithmetic 754-2008. 70 p. doi: 10.1109/IEEESTD.2008.4610935.
15. Fant, K. M. 2005. *Logically determined design: Clockless system design with NULL convention logic*. New York, NY: John Wiley. 292 p. doi: 10.1002/0471702897.
16. Edwards, D., A. Bardsley, L. Jani, L. Plana, and W. Toms. 2006. *Balsa: A tutorial guide*. Manchester. 157 p. Available at: <https://apt.cs.manchester.ac.uk/ftp/pub/amulet/balsa/3.5/BalsaManual3.5.pdf> (accessed March 6, 2025).
17. Tailor, R. A., and R. B. Reese. 2019. UNCLE — Unified NCL Environment — an NCL design tool. *Asynchronous circuit applications*. Institute of Engineering and Technology. 293–307. doi: 10.1049/PBCS061E_ch14.
18. Sokolov, I. A., Yu. A. Stepchenkov, S. G. Bobkov, V. N. Zakharov, Yu. G. Diachenko, Yu. V. Rozhdestvenskiy, and A. V. Surkov. 2014. Bazis realizatsii super-EVM eksaflopsnogo klassa [Implementation basis of exaflops class supercomputer]. *Informatika i ee Primeneniya — Inform. Appl.* 8(1):45–70. doi: 10.14357/19922264140106. EDN: RYYFGD.
19. Rogdestvenskiy, Yu. V., N. V. Morozov, Yu. A. Stepchenkov, and A. V. Rozhdestvenskene. 2006. Universal'naya podsistema analiza samosinkhronnykh skhem [Universal subsystem for self-timed circuits analysis]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 16(1):463–475. EDN: KZUWON.
20. Rogdestvenski, Yu. V., N. V. Morozov, and A. V. Rozhdestvenskene. 2010. Podsystema sobytiynogo analiza samosinkhronnykh skhem ASPEKT [ASPECT: A suite of self-timed event-driven analysis]. *Problemy razrabotki perspektivnykh mikro- i nanoelektronnykh sistem* [Problems of the perspective micro- and nanoelectronic systems development]. Moscow: IPPM RAN. 26–31.
21. Plekhanov, L. P., A. N. Denisov, Yu. G. Diachenko, Yu. A. Stepchenkov, D. I. Mamonov, and D. Yu. Stepchenkov. 2019. Sintez samosinkhronnykh skhem v bazise BMK [Self-timed circuit synthesis in gate array basis]. *5-ya Mezhdunar. nauchn. konf. "Elektronnaya komponentnaya baza i mikroelektronnye moduli": Sb. tezisev* [5th Scientific Conference (International) "Electronic Component Base and Microelectronic Modules": Collection of Abstracts]. Moscow: Tekhnosfera. 450–454.
22. Stepchenkov, Y. A., V. N. Zakharov, Y. G. Diachenko, N. V. Morozov, and D. Y. Stepchenkov. 2015. Cell library for speed-independent VLSI. *IEEE East-West Design and Test Symposium Proceedings*. IEEE. 137–140. doi: 10.1109/EWDTS.2015.7493111.
23. Stepchenkov, Yu. A., A. N. Denisov, Yu. G. Diachenko, et al. 2017. *Biblioteka funktsional'nykh yacheek dlya proektirovaniya samosinkhronnykh poluzakaznykh BMK*

- mikroskhem seriy 5503/5507* [Library of functional cells for designing self-timed semicustom chips of the 5503 and 5597 series]. Moscow: Tekhnosfera. Vol. 4. 376 p.
24. Sokolov, I. A., Yu. A. Stepchenkov, and Yu. G. Dyachenko. 2012. Self-timed RS-trigger with the enhanced noise immunity. U.S. Patent No. 8232825. 31 p.
 25. Sokolov, I. A., Yu. A. Stepchenkov, and Yu. G. Dyachenko. 2012. Self-timed trigger with single-rail data input. U.S. Patent No. 8324938. 34 p.
 26. Zakharov, V. N., Yu. A. Stepchenkov, Yu. G. Diachenko, L. P. Plekhanov, N. V. Morozov, and D. Yu. Diachenko. 2025. Samosinkhronnyy dvukhtaktnyy trigger s parafaznymi vkhodnymi i vykhodnymi signalami s nulevym speyserom [Self-timed push-pull trigger with paraphase input and output signals with zero spacer]. Patent RF No. 2835382.
 27. Stepchenkov, Yu. A., V. S. Petrukhin, and Yu. G. Diachenko. 2006. Opyt razrabotki samosinkhronnogo yadra mikrokontrollera na bazovom matrichnom kristalle [Experience in self-timed microcontroller core design on basic gate-array]. *Nano- i mikrosistemnaya tekhnika* [Nano- and Microsystems Technology] 5:29–36. EDN: IAGLLN.
 28. Stepchenkov, Yu. A., Yu. G. Diachenko, Yu. V. Rogdestvenski, D. Yu. Diachenko and Y. I. Shikunov. 2020. Self-timed multiply-add-subtract unit alternates. *Conference of Russian Young Researchers in Electrical and Electronic Engineering Proceedings*. IEEE. 1864–1868. doi: 10.1109/EIConRus49466.2020.9039039.
 29. Stepchenkov, Yu. A., Yu. G. Diachenko, Yu. V. Rogdestvenski, et al. 2019. Optimizatsiya indikatsii mnogorazryadnykh samosinkhronnykh skhem [Indication optimization in multibit self-timed circuits]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(4):14–27. doi: 10.14357/08696527190402. EDN: XFVKVZ.
 30. Stepchenkov, Yu. A., Yu. G. Diachenko, D. Yu. Stepchenkov, et al. 2023. Mul'tipleksiruemyy samosinkhronnyy konveyer [Multiplexed selftimed pipeline]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 33(2):4–12. doi: 10.14357/08696527230201. EDN: FLVEWP.
 31. Stepchenkov, Yu. A., Yu. G. Diachenko, N. V. Morozov, et al. 2023. Optimizatsiya samosinkhronnogo konveyera [Self-timed pipeline optimization]. *Sistemy vysokoy dostupnosti* [Highly Available Systems] 19(1):5–13. doi: 10.18127/j20729472-202301-01. EDN: IRTFWY.
 32. Stepchenkov, Yu. A., A. N. Kamenskiy, Yu. G. Diachenko, Yu. V. Rogdestvenski, and D. Yu. Diachenko. 2020. Improvement of the natural self-timed circuit tolerance to short-term soft errors. *Advances Science Technology Engineering Systems J.* 5(2):44–56. doi: 10.25046/aj050206.
 33. Sokolov, I., Yu. Stepchenkov, Yu. Diachenko, and D. Khilko. 2025. Mathematical models of critical soft error in synchronous and self-timed pipeline. *Mathematics* 13(5):695. 15 p. doi: 10.3390/math13050695.
 34. Stepchenkov, Yu. A., Yu. G. Diachenko, G. S. Appolonov, et al. 2025. Self-timed pipeline protection against critical fault. *Conference of Russian Young Researchers in Electrical and Electronic Engineering Proceedings*. IEEE. 179–182.
 35. Sokolov, I. A., Yu. A. Stepchenkov, Yu. G. Diachenko, N. V. Morozov, D. Yu. Stepchenkov, and D. Yu. Diachenko. 2022. Analiz sboustoychivosti samosinkhronnogo konveyera [Self-timed pipeline's soft error tolerance analysis]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 32(4):4–13. doi: 10.14357/08696527220401. EDN: NXADJN.

36. Plekhanov, L. P. 22.08.2007. Programma sinteza kombinatsionnykh skhem na zadannoy biblioteke elementov SINTABIB [Program for synthesis of combinational circuits on a given library of elements SINTABIB]. Certificate on official registration of the computer program No. 2007613665.
37. Morozov, N. V., Yu. G. Diachenko, D. Yu. Stepchenkov, and A. V. Rozhdestvenskene. 19.11.2021. Sistema kharakterizatsii samosinkhronnykh elementov SAKhIB. Versiya 4 [SAHIB: Self-timed cell characterization system, version 2]. Certificate on official registration of the computer program No. 2021668787.
38. Zatsarinnyy, A. A., Yu. A. Stepchenkov, Yu. G. D'yachenko, *et al.* 2023. Avtomatizatsiya sinteza samosinkhronnykh skhem [Selftimed circuits design automation]. *Sistemy vysokoy dostupnosti* [Highly Available Systems] 19(3):48–56. doi: 10.18127/j20729472-202303-04. EDN: IAPVXR.

Received December 27, 2024

Accepted February 15, 2025

Contributors

Zakharov Victor N. (b. 1948) — Doctor of Science in technology, associate professor, scientific secretary, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; vzakharov@ipiran.ru

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

Diachenko Yuri G. (b. 1958) — Candidate of Science (PhD) in technology, senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; diaura@mail.ru

Morozov Nikolai V. (b. 1956) — senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; NMorozov@ipiran.ru

Plekhanov Leonid P. (b. 1943) — Candidate of Science (PhD) in technology, senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; lplekhanov@inbox.ru

Stepchenkov Dmitri Yu. (b. 1973) — senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; stepchenkov@mail.ru

ПРОЦЕССОРЫ ДАННЫХ КАК НОВЫЙ КОМПОНЕНТ ИНФРАСТРУКТУРЫ ДАТА-ЦЕНТРОВ

*В. Б. Егоров*¹

Аннотация: Традиционно обрабатывающие возможности дата-центров реализовывались процессорами трех типов: центральными (central processing units, CPU), графическими (graphics processing units, GPU) и сетевыми (network processing units, NPU). В последнее время к этой триаде присоединились процессоры данных (data processing units, DPU), иногда также называемые инфраструктурными процессорами. Наиболее часто декларируемое в рекламе назначение DPU — освобождение CPU и GPU от коммуникационных нагрузок и ряда рутинных работ по хранению, перемещению и обработке некоторых типов данных в масштабах дата-центра, что предполагает взаимодействие DPU с CPU и GPU, а также частичную или полную замену ими NPU. В статье делается попытка прояснить сущность DPU, уточняется их место и роль в инфраструктуре дата-центров, рассматривается связь DPU с GPU и NPU, выясняется отношение DPU к виртуализации сетевых функций, обозначаются имеющиеся трудности их использования, дается предварительная оценка перспектив дальнейшего их развития и применения.

Ключевые слова: виртуализация сетевых функций; графический процессор; процессор данных; сетевой процессор; центральный процессор

DOI: 10.14357/08696527250109

EDN: VBYUTE

1 Введение

Традиционно основные вычислительные возможности дата-центров реализовывались универсальными центральными процессорами (CPU), к которым еще в прошлом веке присоединились специализированные графические (GPU) и сетевые процессоры (NPU). С тех пор пара CPU + GPU стала неразлучной в самых разных реализациях, включая интеграцию в одном приборе — так называемом «ускоренном» процессоре (accelerated processing unit, APU) — и в самом широком спектре устройств от смартфонов до серверов и суперкомпьютеров. При этом NPU в конечном счете не нашли массового применения, в настоящее время их выпуск как самостоятельных продуктов практически прекращен [1], а на освободившуюся нишу специализированного сетевого интеллекта обозначил претензии новый класс интегрированных приборов — процессоры данных (DPU)².

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

²Следует обратить внимание, что исчезновение сетевых процессоров из номенклатуры инфраструктурных устройств дата-центров освободило не только нишу сетевого интеллекта, но и аббревиатуру NPU, которую немедленно присвоили себе так называемые «нейронные» процессоры (neural processing units). Здесь эта аббревиатура относится исключительно к сетевым процессорам.

Несмотря на явно неудачное название, так как в конечном счете все процессоры, в том числе CPU, GPU и NPU, так или иначе обрабатывают те или иные данные, название успело прижиться, и аббревиатура DPU все чаще встречается в рекламе. Альтернативу этому неудачному названию предложила компания Intel, которая выпустила в содружестве с Google Cloud прибор (кодовое обозначение Mount Evans) с заявленными функциями, в целом типичными для DPU, но презентовала его в качестве инфраструктурного процессора (infrastructure processing unit) [2, 3]. Однако ощутимой поддержки эта альтернатива не получила, может быть, потому, что и она не проясняет сущность и роль нового класса процессоров, а лишь декларирует их принадлежность к инфраструктуре дата-центров.

В настоящей статье делается попытка прояснить сущность DPU, обозначить их место в инфраструктуре дата-центров, выявить взаимосвязь с CPU, GPU и NPU и их отношение к виртуализации сетевых функций (network functions virtualization, NFV), а также предварительно оценить перспективы их развития и использования.

2 Назначение процессоров данных и их роль в дата-центре

Декларируемое в большинстве презентаций и публикаций назначение DPU — снять сетевую и коммуникационную нагрузку с CPU и GPU, взяв на себя рутинные работы по хранению, перемещению и обработке некоторых типов данных в масштабах дата-центра, поскольку такого рода работы могут отнимать значительную часть процессорных ресурсов [4–8]. Например, в современных виртуализированных средах обработка транзакций виртуального коммутатора Open vSwitch может занимать более 30% рабочего времени CPU [9]. Типичные возлагаемые на DPU обязанности могут включать собственно передачу данных (data transfer) с возможностью их шифрования и защиты, сжатие данных (data compression), сокращение занимаемых ими объемов памяти (data reduction), а также предоставление дополнительных возможностей разного рода анализа и статистики [5].

Реклама DPU обещает, что в ориентированной на данные (data-centric) инфраструктуре дата-центров они смогут заметно улучшить производительность серверов и существенно повысить эффективность хранилищ данных, включая сетевые хранилища (network attached storage) и сети хранения данных (storage-area networks). Предполагается также, что DPU смогут с успехом применяться в «облачных» дата-центрах и суперкомпьютерах, в частности при решении задач машинного обучения (machine learning, ML) и искусственного интеллекта (artificial intelligence, AI) [5]. При этом эффект от использования DPU ожидается двояким: во-первых, благодаря ускорению решения типовых инфраструктурных задач дата-центра в DPU, более производительных на этих задачах за счет специализации аппаратуры и программного обеспечения (ПО); во-вторых, вследствие освобождения от этих задач CPU и GPU, которые получают возможность обслуживать большее число приложений.

На сегодняшний день основной, если не сказать единственной, областью применения DPU стали интеллектуальные сетевые карты [4]. Сетевые карты (network interface cards, NIC), традиционно служащие для подключения серверов к сетям передачи данных, обычно строятся на основе специализированных интегральных схем, и их жестко заданную разработчиком функциональность невозможно изменить. В отличие от обычных NIC, интеллектуальные сетевые карты smartNIC позволяют загружать в них дополнительное, в том числе пользовательское, ПО, которое расширяет функциональные возможности карт, например добавляя глубокую инспекцию пакетов (deep packet inspection, DPI) на сетевых интерфейсах или поддержку протоколов доступа к твердотельным накопителям NVMe и NVMe-oF по интерфейсу ввода-вывода [9, 10]. Таким образом, внедрение DPU в smartNIC расширяет функциональные возможности последних и делает их более гибкими и лучше адаптируемыми к меняющимся условиям современных быстро развивающихся дата-центров.

3 Процессоры данных и виртуализация сетевых функций

В качестве одного из положительных следствий внедрения DPU предполагается возможность снизить общие затраты на серверное оборудование дата-центра, поскольку DPU должны освободить CPU от целого ряда несвойственных им задач и тем самым позволить обходиться меньшим числом серверов, не снижая общую производительность дата-центра [8]. Тезис этот не очевиден и требует более конкретной оценки затрат в масштабе дата-центра не только на CPU, которых, может быть, потребуется меньше, но и на DPU с их специфическим ПО как дополнительную статью расходов.

Кроме того, этот неочевидный тезис формально входит в противоречие с концепцией NFV, одна из основных целей которой состоит в замене специализированного сетевого оборудования (коммутаторов, балансировщиков нагрузки, межсетевых экранов, систем выявления и отражения вторжений и т. п.) серийными (off-the-shelf) универсальными серверами, на которых соответствующим ПО эмулируются традиционные сетевые устройства [11]. При этом помимо повышения гибкости и адаптируемости инфраструктуры дата-центра также предполагается снижение стоимости оборудования, поскольку массово выпускаемые универсальные серверы оказываются, как правило, дешевле совокупно заменяемой специализированной сетевой аппаратуры, к которой формально можно отнести и DPU.

Разрешение этого бросающегося в глаза конфликта DPU с NFV, вероятно, будет зависеть от практических оценок эффективности DPU и их относительной стоимости, которая будет определяться в числе прочего их востребованностью и, как следствие, массовостью выпуска. Согласно прогнозу Allied Market Research [8], мировой рынок DPU будет расти в среднем на 26,9% в год. Если этот оптимистичный прогноз сбудется, DPU со временем должны превратиться в массовый рыночный продукт с объемами выпуска, соизмеримыми с объема-

ми выпуска CPU и GPU. Однако, как показывает пример предшественников DPU — интегрированных NPU, далеко не все благоприятные прогнозы сбываются.

4 Процессоры данных и сетевые процессоры

Типичный DPU представляет собой интегрированный прибор, который может включать [6, 8]:

- универсальный процессор, как правило, многоядерный, в большинстве случаев с архитектурой ARM;
- множество специализированных программируемых процессоров с оригинальными системами инструкций, в частности ориентированными на задачи управления сетями и хранилищами данных;
- разного рода аппаратные ускорители;
- высокоскоростное подключение к сети Ethernet, в некоторых случаях также к сети InfiniBand;
- интерфейс ввода-вывода с высокой пропускной способностью, как правило, многодорожечный PCI Express.

Нетрудно видеть, что по своей «начинке» DPU практически не отличаются от интегрированных NPU [1]. Неудивительно, что некоторые имеющиеся на рынке DPU на поверку оказываются просто «перелицованными» NPU. Хороший тому пример — DPU компании Marvell, которые представляют собой реинкарнацию интегрированных NPU семейства OSTEON, причем наряду с относительно новыми линейками приборов с процессорными ядрами ARM в качестве DPU предлагаются и более старые линейки NPU с ядрами архитектуры MIPS [12].

Таким образом, DPU в значительной степени наследуют от интегрированных NPU не только область применения, но и архитектурные принципы построения, вследствие чего им потенциально грозят те же трудности, с которыми встретились и которые не смогли преодолеть NPU. В первую очередь это относится к разнообразию конкретных архитектурных решений специализированных процессорных ядер с вытекающими из этого разнообразия сложностями программирования. Не лучше ситуация и с аппаратными ускорителями, номенклатура, функциональность и реализация которых никак не регламентированы и в каждом конкретном DPU произвольно определяются разработчиком.

Программируемость заявляется как одно из важнейших требований к DPU. Однако у программируемости имеется обратная сторона: уникальность каждого отдельно взятого дата-центра и совокупности выполняемых в нем работ превращает программирование DPU из возможности в необходимость. Между тем, в отличие от CPU, для DPU программировать приходится не приложения на удобных языках под стандартные операционные среды с использованием апробированных средств разработки, а специфические инфраструктурные задачи,

разнящиеся в разных дата-центрах, к тому же в средах разработки, специфичных для конкретных DPU. Аналогичная проблема программирования NPU так и не была решена и стала камнем преткновения на их пути к массовому применению. Попутно можно отметить, что во многом похожие трудности создания уникального ПО для каждого конкретного дата-центра уже не один год тормозят массовое внедрение программно-определяемых сетей (software-defined networking, SDN) [13].

5 Процессоры данных и графические процессоры

Недавно на совсем еще молодой рынок DPU неожиданно дружно вышли компании NVIDIA и AMD — безусловные лидеры и главные конкуренты в области высокопроизводительных GPU. И этот факт достоин внимания.

NVIDIA предлагает включающую DPU smartNIC BlueField [14], которая была приобретена вместе с ее разработчиком, компанией Mellanox (2020 г.). Ранее Mellanox специализировалась на сетевой аппаратуре и интегрированных NPU, и DPU семейства BlueField стали продолжателями интегрированных NPU семейства Tile-Gx/Mx компании EZchip, ранее поглощенной Mellanox (2016 г.). Таким образом, DPU BlueField оказываются прямыми наследниками нескольких поколений NPU. И эта наследственность неслучайна. В конечном счете, как было отмечено выше, NPU и DPU не только решают схожие задачи, но и используют одни и те же средства их решения.

По замыслу компании NVIDIA, ее DPU (самый последний в семействе — BlueField-3 [15]) должны дополнить фирменные сверхмощные универсальные CPU (пилотный прибор — суперчип Grace CPU [16]) и широко известные GPU с тензорными ядрами (последние архитектуры — Норрег и Blackwell [17]), став третьей опорой, обеспечивающей устойчивость конструкции инфраструктур, ориентированных в первую очередь на высокопроизводительные вычисления (high-performance computing, HPC), а также задачи ML и AI. В целом сферу применения своих DPU NVIDIA хотела бы видеть предельно широкой [9], включающей:

- виртуальные и аппаратные «облачные» среды;
- хранилища NVMe в виртуальных машинах;
- NFV;
- DPI;
- микросерверы для граничных вычислений.

Вслед за NVIDIA тем же путем пошел ее главный конкурент в области GPU, компания AMD, которая приобрела компанию Pensando Systems (2022 г.) вместе с ее разработкой — smartNIC Pensando DSC-25/100, позиционируемой как карта распределенных сервисов (distributed services card). На этом новом для

себя направлении деятельности AMD предлагает DPU Giglio и Salina [18], сохраняющие совместимость с процессорами smartNIC семейства Pensando (Capri, Elba) и при этом, согласно рекламе AMD, «обеспечивающие невероятную энергоэффективность и в то же время значительно увеличивающие пропускную способность конвейеров P4» [19].

Тот примечательный факт, что два мировых лидера в области GPU практически одновременно вышли на рынок DPU, причем в обоих случаях приобретением компаний, специализировавшихся в области smartNIC, дает основания ожидать как дальнейшего развития DPU в составе smartNIC, так и, гипотетически, возможного выхода DPU за рамки сетевых карт, объединения их возможностей с возможностями GPU. В частности, связка GPU + DPU может оказаться эффективной в решении задач НРС/ML/AI, а также в более тесном сближении вычислений с хранением данных, движении в сторону «вычислительного хранилища» (computational storage) [20].

6 Заключение

Процессоры данных как новый тип обрабатывающего устройства декларирован в качестве средства освобождения традиционных обрабатывающих устройств, CPU и GPU, от рутинных не свойственных им задач по обслуживанию инфраструктуры дата-центров. В результате эффективность дата-центров повышается благодаря специализации DPU под эти задачи и вследствие лучшего использования CPU и GPU. В значительной степени DPU стали наследниками интегрированных NPU как по кругу решаемых задач, так и по средствам их решения. Как следствие, DPU частично унаследовали и проблемы NPU, в частности в отношении программируемости.

На сегодняшний день, несмотря на активную рекламу и благоприятные прогнозы, можно отметить лишь отдельные, хотя и интересные сами по себе, попытки использования DPU, пока что исключительно в составе smartNIC семейства BlueField, в качестве компонента систем, оптимизированных для приложений с интенсивной работой в сети [21], в том числе со многими арендаторами (multitenant) [22], и «облачных» сред бессерверных вычислений (serverless computing) [23]. Поэтому вряд ли сегодня можно с уверенностью прогнозировать широкое распространение DPU в качестве самостоятельного инфраструктурного компонента. На данный момент они остаются локальным интеллектом в рамках smartNIC, причем конкретные DPU связаны с конкретными smartNIC общим разработчиком. В этом плане показательно, что у компании NVIDIA бренд BlueField относится как к DPU, так и к построенным на их основе smartNIC (SuperNIC), т.е. фактически DPU не рассматривается компанией как самостоятельный продукт [15, 24]. Пока что для DPU как отдельного рыночного продукта существует риск повторить незавидную судьбу NPU и остаться «все-го лишь внутрифирменными инструментами интеллектуализации ограниченного

числа специфических продуктов в узком кругу крупных микроэлектронных фирм» [1].

Возможно, ситуацию могло бы изменить появление в следующем поколении вычислительной техники некоего «отраслевого стандарта» [8], который отвяжет производителей smartNIC от разработчиков DPU. Кроме того, вероятным представляется появление стандартов де-факто от компании NVIDIA, взявшей курс на оригинальные, не скованные никакими стандартами сверхвысокопроизводительные решения в области HPC/ML/AI, основываясь на фирменных тесно интегрированных APU. Также одним из шагов к таким стандартам де-факто могут стать работы NVIDIA в партнерстве с компанией Pluribus Networks по интеграции DPU с сетевой операционной системой Pluribus NetVisor [25] в качестве, например, основы встраиваемых в smartNIC SDN-контроллеров.

Литература

1. *Егоров В. Б.* Эволюция сетевых процессоров // Системы и средства информатики, 2021. Т. 31. № 1. С. 111–121. doi: 10.14357/08696527210109. EDN: CTKUMR.
2. *Kumar R.* Intel E2000 is the new Intel Mount Evans DPU IPU brand. — STH, 2022. <https://www.servethehome.com/intel-e2000-is-the-new-intel-mount-evans-dpu-ipu-brand/#:~:text=The%20Intel%20Mount%20Evans%20IPU,Data%20Processing%20Unit%20Quick%20Primer.>
3. *Salamone S.* IPU chip offloads networking and some security tasks from CPUs. — Network Computing, 2022. <https://www.networkcomputing.com/data-centers/ipu-chip-offloads-networking-and-some-security-tasks-cpus.>
4. *Deierling K.* What is a DPU? — NVIDIA, 2020. <https://blogs.nvidia.com/blog/2020/05/20/whats-a-dpu-data-processing-unit.>
5. *Roundy J.* How do CPU, GPU and DPU differ from one another? — TechTarget, 2021. <https://www.techtarget.com/searchdatacenter/tip/How-do-CPU-GPU-and-DPU-differ-from-one-another.>
6. *Trick C.* What is a DPU (Data processing unit)? — Trenton Systems, 2022. <https://www.trentonsystems.com/blog/what-is-a-dpu.>
7. *Kerravala Z.* Understanding the power benefits of data processing units. — Network Computing, 2022. <https://www.networkcomputing.com/data-centers/understanding-power-benefits-data-processing-units.>
8. *Eileen A.* Data processing units: What are DPUs and why do you want them? // Data Center Knowledge, 2023. <https://www.datacenterknowledge.com/data-center-faqs/data-processing-units-what-are-dpus-and-why-do-you-want-them#close-modal.>
9. *Турекс.* Как устроены DPU, сопроцессоры для обработки данных. — Академия Selectel, 2020. [https://selectel.ru/blog/dpu/.](https://selectel.ru/blog/dpu/)
10. BCM58800. High-performance datacenter SoC with integrated NetXtreme Ethernet controller. — Broadcom, 2017. <https://docs.broadcom.com/doc/58800-PB100.>
11. *Егоров В. Б.* Некоторые неявные вопросы сетевой виртуализации // Системы и средства информатики, 2017. Т. 27. № 3. С. 88–98. doi: 10.14357/08696527170308. EDN: ZQLXVN.

12. Data Processing Units (DPU). — Marvell, 2023. <https://www.marvell.com/products/data-processing-units.html>.
13. Егоров В. Б. Некоторые вопросы практической реализации концепции SDN // Системы и средства информатики, 2016. Т. 26. № 1. С. 109–120. doi: 10.14357/08696527160108. EDN: VZBFNT.
14. NVIDIA BlueField-2 DPU. Data center infrastructure on a chip. — NVIDIA, 2021. <https://www.nvidia.com/content/dam/en-zz/Solutions/Data-Center/documents/datasheet-nvidia-bluefield-2-dpu.pdf>.
15. NVIDIA BlueField networking platform. — NVIDIA, 2023. <https://www.nvidia.com/en-us/networking/products/data-processing-unit/>.
16. NVIDIA Grace CPU superchip. — NVIDIA, 2023. <https://www.nvidia.com/en-us/data-center/grace-cpu-superchip/>.
17. NVIDIA GH200 Grace Hopper superchip. — NVIDIA, 2023. <https://www.nvidia.com/en-us/data-center/grace-hopper-superchip/>.
18. Степин А. AMD готовит новые DPU Pensando: Giglio и Salina // ServerNews, 2022. <https://servernews.ru/1073294>; <https://dzen.ru/a/Yw9dPnlTczKSXvEy>.
19. AMD Pensando networking. — AMD, 2024. <https://www.amd.com/en/products/accelerators/pensando.html>.
20. Davis R. GPU + DPU for computational storage. — SNIA, 2022. <https://www.snia.org/sites/default/files/PM-Summit/2022/PMCS22-Davis-GPU-and-DPU-in-Computational-Storage.pdf>.
21. Huang J., Lou J., Sun Y., Wang T., Lee E. K., Kim N. S. analyzing energy efficiency of a server with a SmartNIC under SLO constraints // Symposium (International) on Performance Analysis of Systems and Software Proceedings. — Piscataway, NJ, USA: IEEE, 2023. P. 334–336. doi: 10.1109/ISPASS57527.2023.00044.
22. Huang Z., Tan Y., Zhu Y., Tan H., Li K. MTDA: Efficient and fair DPU offloading method for multiple tenants // IEEE T. Serv. Comput., 2024. Vol. 17. No. 6. P. 3971–3984. doi: 10.1109/TSC.2024.3433588.
23. Du D., Liu Q., Jiang X. Q. Serverless computing on heterogeneous computers // 27th Conference (International) on Architectural Support for Programming Languages and Operating Systems Proceedings. — New York, NY, USA: ACM, 2022. P. 797–813. doi: 10.1145/3503222.3507732.
24. Карасев С. NVIDIA представила сетевой ускоритель SuperNIC для гипермасштабируемых ИИ-нагрузок // ServerNews, 2023. <https://servernews.ru/1096340>.
25. Raynovich S. Networking could be NVIDIA's next big game // Forbes, 2022. <https://www.forbes.com/sites/rscottraynovich/2022/03/28/network-could-be-nvidias-next-big-game/?sh=7c06bc201bcd>.

Поступила в редакцию 19.07.2024

Принята к публикации 15.02.2025

DATA PROCESSING UNITS AS A NEW COMPONENT OF THE DATA CENTER INFRASTRUCTURE

V. B. Egorov

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

Abstract: The data center processing capabilities have been traditionally implemented by processors of three kinds — central processing units (CPU), graphics processing units (GPU), and network processing units (NPU). Recently, data processing units (DPU), sometimes called also infrastructure processors, have joined this triad. The most frequently advertised purpose of the DPU is to free CPU and GPU from communication burdens and a number of routine tasks for storing, moving, and processing certain data types within a data center that assumes interaction of DPU with CPU and GPU as well as partial or complete replacement of the NPU. The article attempts to clarify the essence of DPU, specifies their place and role in the data center infrastructure, examines the connection of DPU with GPU and NPU, finds out the relation to network functions virtualization, denotes difficulties of using the DPU, and gives a preliminary assessment of the prospects for their further development and application.

Keywords: central processing unit; data processing unit; graphics processing unit; network functions virtualization; network processing unit

DOI: 10.14357/08696527250109

EDN: VBYUTE

References

1. Egorov, V. B. 2021. Evolyutsiya setevykh protsessorov [Evolution of network processors]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 31(1):111–121. doi: 10.14357/08696527210109. EDN: CTKUMR.
2. Kumar, R. 2022. Intel E2000 is the new Intel Mount Evans DPU IPU brand. *STH*. Available at: <https://www.servethehome.com/intel-e2000-is-the-new-intel-mount-evans-dpu-ipu-brand/#:~:text=The%20Intel%20Mount%20Evans%20IPU,Data%20Processing%20Unit%20Quick%20Primer> (accessed March 10, 2025).
3. Salamone, S. 2022. IPU chip offloads networking and some security tasks from CPUs. *Network Computing*. Available at: <https://www.networkcomputing.com/data-centers/ipu-chip-offloads-networking-and-some-security-tasks-cpus> (accessed March 10, 2025).
4. Deierling, K. 2020. What is a DPU? NVIDIA. Available at: <https://blogs.nvidia.com/blog/2020/05/20/whats-a-dpu-data-processing-unit> (accessed March 10, 2025).
5. Roundy, J. 2021. How do CPU, GPU and DPU differ from one another? *TechTarget*. Available at: <https://www.techtarget.com/searchdatacenter/tip/How-do-CPU-GPU-and-DPU-differ-from-one-another> (accessed March 10, 2025).
6. Trick, C. 2022. What is a DPU (Data processing unit)? *Trenton Systems*. Available at: <https://www.trentonsystems.com/blog/what-is-a-dpu> (accessed March 10, 2025).

7. Kerravala, Z. 2022. Understanding the power benefits of data processing units. Network Computing. Available at: <https://www.networkcomputing.com/data-centers/understanding-power-benefits-data-processing-units> (accessed March 10, 2025).
8. Eileen, A. 2023. Data processing units: What are DPUs and why do you want them? Data Center Knowledge. Available at: <https://www.datacenterknowledge.com/data-center-faqs/data-processing-units-what-are-dpus-and-why-do-you-want-them#close-modal> (accessed March 10, 2025).
9. Kak ustroeny DPU, soprotsessory dlya obrabotki dannykh [How DPUs and coprocessors for data processing work]. 2020. Selectel Academy. Available at: <https://selectel.ru/blog/dpu/> (accessed March 10, 2025).
10. BCM58800. High-performance datacenter SoC with integrated NetXtreme Ethernet controller. 2017. Broadcom. Available at: <https://docs.broadcom.com/doc/58800-PB100> (accessed March 10, 2025).
11. Egorov, V. 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. doi: 10.14357/08696527170308. EDN: ZQLXVN.
12. Data processing units (DPU). 2023. Marvell. Available at: <https://www.marvell.com/products/data-processing-units.html> (accessed March 10, 2025).
13. 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. doi: 10.14357/08696527160108. EDN: VZBFNT.
14. NVIDIA BlueField-2 DPU. 2021. Data center infrastructure on a chip. NVIDIA. Available at: <https://www.nvidia.com/content/dam/en-zz/Solutions/Data-Center/documents/datasheet-nvidia-bluefield-2-dpu.pdf> (accessed March 10, 2025).
15. NVIDIA BlueField networking platform. 2023. NVIDIA. Available at: <https://www.nvidia.com/en-us/networking/products/data-processing-unit/> (accessed March 10, 2025).
16. NVIDIA Grace CPU superchip. 2023. NVIDIA. Available at: <https://www.nvidia.com/en-us/data-center/grace-cpu-superchip/> (accessed March 10, 2025).
17. NVIDIA GH200 Grace Hopper superchip. 2023. NVIDIA. Available at: <https://www.nvidia.com/en-us/data-center/grace-hopper-superchip/> (accessed March 10, 2025).
18. Stepin, A. 2022. AMD gotovit novye DPU Pensando: Giglio i Salina [AMD is preparing new DPUs Pensando: Giglio and Salina]. *ServerNews*. Available at: <https://servernews.ru/1073294>; <https://dzen.ru/a/Yw9dPnIczKSXvEy> (accessed March 10, 2025).
19. AMD Pensando networking. 2024. AMD. Available at: <https://www.amd.com/en/products/accelerators/pensando.html> (accessed March 10, 2025).
20. Davis, R. 2022. GPU+DPU for computational storage. SNIA. Available at: <https://www.snia.org/sites/default/files/PM-Summit/2022/PMCS22-Davis-GPU-and-DPU-in-Computational-Storage.pdf> (accessed March 10, 2025).
21. Huang, J., J. Lou, Y. Sun, T. Wang, E. K. Lee, and N. S. Kim. 2023. Analyzing energy efficiency of a server with a SmartNIC under SLO constraints. *Symposium (Internationa-*

- tional) on Performance Analysis of Systems and Software Proceedings. Piscataway, NJ: IEEE. 334–336. doi: 10.1109/ISPASS57527.2023.00044.
22. Huang, Z., Y. Tan, Y. Zhu, H. Tan, and K. Li. 2024. MTDA: Efficient and fair DPU offloading method for multiple tenants. *IEEE T. Serv. Comput.* 17(6):3971–3984. doi: 10.1109/TSC.2024.3433588.
23. Du, D., Q. Liu, and X. Q. Jiang. 2022. Serverless computing on heterogeneous computers. *Conference (International) on Architectural Support for Programming Languages and Operating Systems Proceedings*. New York, NY: ACM. 797–813. doi: 10.1145/3503222.3507732.
24. Karasev, S. 2023. NVIDIA predstavila setevoy uskoritel' SuperNIC dlya giper-masshtabiruemykh II-nagruzok [NVIDIA introduced SuperNIC network accelerator for hyper-scale AI workloads]. *ServerNews*. Available at: <https://servernews.ru/1096340> (accessed March 10, 2025).
25. Raynovich, S. 2022. Networking could be NVIDIA's next big game. *Forbes*. Available at: <https://www.forbes.com/sites/rscottraynovich/2022/03/28/network-could-be-nvidias-next-big-game/?sh=7c06bc201bcd> (accessed March 10, 2025).

Received July 19, 2024

Accepted February 15, 2025

Contributor

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

Брюхов Дмитрий Олегович (р. 1971) — кандидат технических наук, старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Гридин Владимир Николаевич (р. 1944) — доктор технических наук, научный руководитель Центра информационных технологий в проектировании Российской академии наук

Грушо Александр Александрович (р. 1946) — доктор физико-математических наук, профессор, главный научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Грушо Николай Александрович (р. 1982) — кандидат физико-математических наук, старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Добровольский Дмитрий Олегович (р. 1953) — доктор филологических наук, главный научный сотрудник Института русского языка Российской академии наук; главный научный сотрудник Института языкознания Российской академии наук; главный научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Дьяченко Юрий Георгиевич (р. 1958) — кандидат технических наук, старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Егоров Владимир Борисович (р. 1948) — кандидат технических наук, ведущий научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Забейхайло Михаил Иванович (р. 1956) — доктор физико-математических наук, профессор, главный научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Захаров Виктор Николаевич (р. 1948) — доктор технических наук, доцент, ученый секретарь Федерального исследовательского центра «Информатика и управление» Российской академии наук

Зацаринный Александр Алексеевич (р. 1951) — доктор технических наук, профессор, главный научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Зацман Игорь Моисеевич (р. 1952) — доктор технических наук, заведующий отделом 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

Колин Константин Константинович (р. 1935) — доктор технических наук, профессор, заслуженный деятель науки РФ, главный научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Листопад Сергей Викторович (р. 1984) — кандидат технических наук, старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

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

Плеханов Леонид Петрович (р. 1943) — кандидат технических наук, старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

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

Салем Басим Раед (р. 1997) — научный сотрудник Центра информационных технологий в проектировании Российской академии наук

Синицын Игорь Николаевич (р. 1940) — доктор технических наук, профессор, заслуженный деятель науки РФ, главный научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук; профессор кафедры «Мехатроника и теоретическая механика» Московского авиационного института (национального исследовательского университета)

Скворцов Николай Алексеевич (р. 1973) — научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Смирнов Дмитрий Владимирович (р. 1984) — бизнес-партнер по ИТ департамента безопасности ПАО «Сбербанк России»

Смирнов Дмитрий Сергеевич (р. 1988) — кандидат экономических наук, заведующий лабораторией Центра информационных технологий в проектировании Российской академии наук

Соколов Игорь Анатольевич (р. 1954) — доктор технических наук, академик РАН, научный руководитель Федерального исследовательского центра «Информатика и управление» Российской академии наук

Солодовников Владимир Игоревич (р. 1977) — кандидат технических наук, директор Центра информационных технологий в проектировании Российской академии наук

Степченков Дмитрий Юрьевич (р. 1973) — старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Степченков Юрий Афанасьевич (р. 1951) — кандидат технических наук, заведующий отделом, ведущий научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Ступников Сергей Александрович (р. 1978) — кандидат технических наук, ведущий научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Тимонина Елена Евгеньевна (р. 1952) — доктор технических наук, профессор, ведущий научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Шоргин Сергей Яковлевич (р. 1952) — доктор физико-математических наук, профессор, главный научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Правила подготовки рукописей статей для публикации в журнале «Системы и средства информатики»

Журнал «Системы и средства информатики» публикует теоретические, обзорные и дискуссионные статьи, посвященные научным исследованиям и разработкам в области информационных технологий.

Журнал издается на русском языке. По специальному решению редколлегии отдельные статьи могут печататься на английском языке.

Тематика журнала охватывает следующие направления:

- информационно-телекоммуникационные системы и средства их построения;
- архитектура и программное обеспечение вычислительных машин, комплексов и сетей;
- методы и средства защиты информации.

1. В журнале печатаются статьи, содержащие результаты, ранее не опубликованные и не предназначенные к одновременной публикации в других изданиях.

Публикация предоставленной автором(ами) рукописи не должна нарушать положений глав 69, 70 раздела VII части IV Гражданского кодекса, которые определяют права на результаты интеллектуальной деятельности и средства индивидуализации, в том числе авторские права, в РФ.

Ответственность за нарушение авторских прав, в случае предъявления претензий к редакции журнала, несут авторы статей.

Направляя рукопись в редакцию, авторы сохраняют свои права на данную рукопись и при этом передают учредителям и редколлегии журнала неисключительные права на издание статьи на русском языке (или на языке статьи, если он отличен от русского) и на перевод ее на английский язык, а также на ее распространение в России и за рубежом. Каждый автор должен представить в редакцию подписанный с его стороны «Лицензионный договор о передаче неисключительных прав на использование произведения», текст которого размещен по адресу <http://www.ipiran.ru/publications/licence.doc>. Этот договор может быть представлен в бумажном (в 2-х экз.) или в электронном виде (отсканированная копия заполненного и подписанного документа).

Если при подготовке статьи авторы использовали инструменты на основе искусственного интеллекта, они обязаны включить эту информацию в текст статьи.

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 — A4; ориентация — книжная; поля (см): внутри — 2,5, снаружи — 1,5, сверху и снизу — 2, от края до нижнего колонтитула — 1,3.

Основной текст: стиль — «Обычный», шрифт — Times New Roman, размер — 14 пунктов, абзацный отступ — 0,5 см, 1,5 интервала, выравнивание — по ширине.

Рекомендуемый объем рукописи — не выше 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 оформления представлены в интернете:

http://www.ipiran.ru/journal/template_iiep_ssi_2024.zip

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 resursosberegayushchie 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 kompaktnyy tor* [Mathematical modeling of the plasma in the compact torus]. Moscow. D.Sc. Diss. 272 p.

Kozhunova, O. S. 2009. *Tekhnologiya razrabotki semanticheskogo slovary 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 kolichstva zhidkostey i gazov s pomoshch'yu standartnykh suzhayushchikh ustroystv* [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. Moscow: Standardinform Publ. 10 p.

Описание патента:

Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. *Sposob orientirovaniya po krenu letatel'nogo apparata s opticheskoy golovkoy samonavedeniya* [The way to orient on the roll of aircraft with optical homing head]. Patent RF No. 2280590.

10. Присланные в редакцию материалы авторам не возвращаются.

11. При отправке файлов по электронной почте просим придерживаться следующих правил:

- указывать в поле subject (тема) название журнала и фамилию автора;
- указывать в тексте письма название статьи, авторов и журнал, в который направляется статья;
- 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.

If authors used artificial intelligence (AI)-based tools in preparing their manuscript, they must include this information in the text of the article.

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/journal/template_jiep_ssi_2024.zip

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 Publ. 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 Publ. 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;
 - the article title, authors' names, and the journal title, whereto the paper is being submitted, in the text of the e-mail;
 - 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