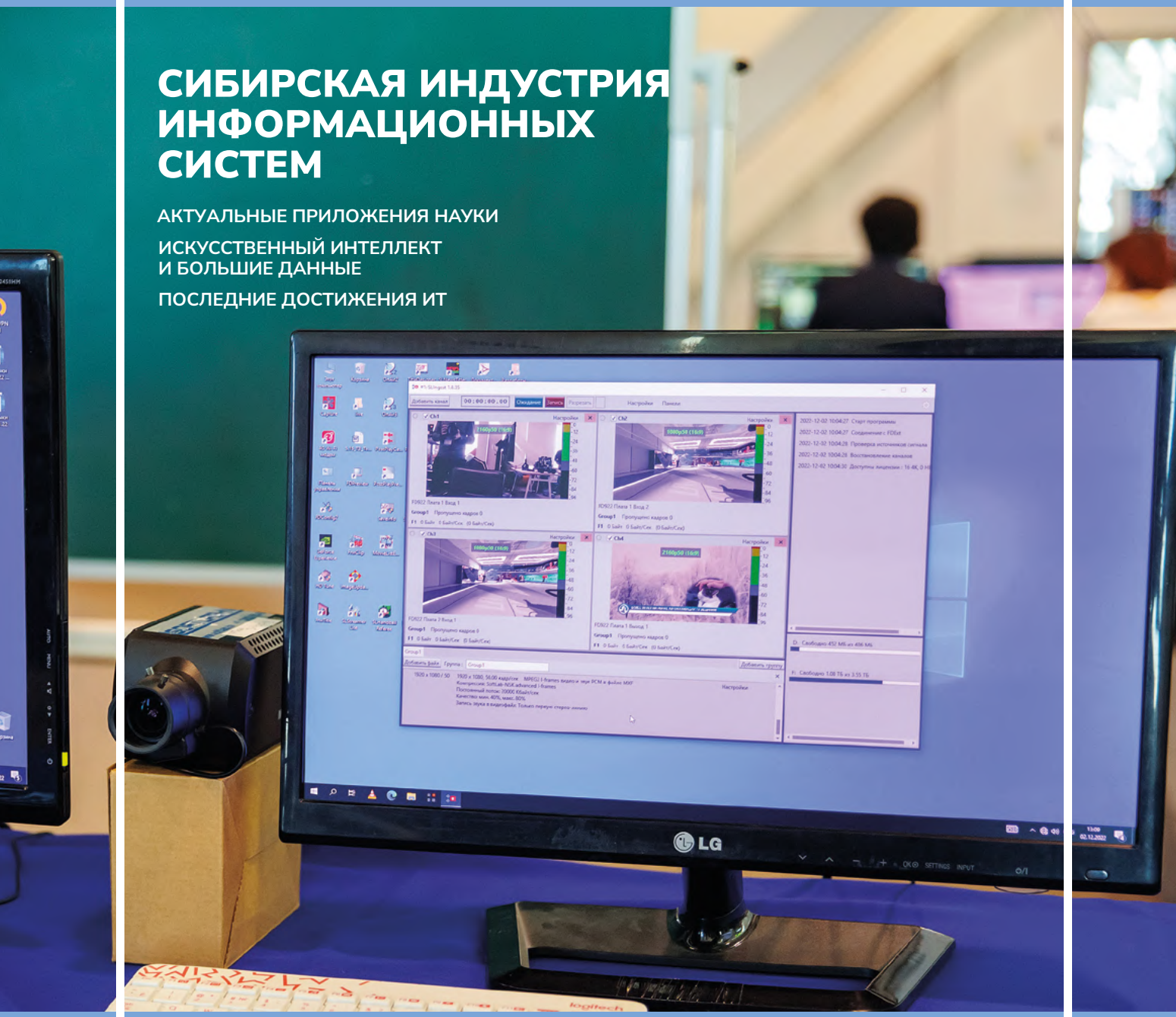
наука

№ 3 (10) | 2023 год
Издание СО РАН



СИБИРСКАЯ ИНДУСТРИЯ ИНФОРМАЦИОННЫХ СИСТЕМ

АКТУАЛЬНЫЕ ПРИЛОЖЕНИЯ НАУКИ

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И БОЛЬШИЕ ДАННЫЕ

ПОСЛЕДНИЕ ДОСТИЖЕНИЯ ИТ

и технологии Сибири

наука
и технологии
Сибири

НАУКА И ТЕХНОЛОГИИ СИБИРИ
Выпуск 10 — Сибирская индустрия информационных систем.
Ноябрь 2023 г.

Учредитель:
Федеральное государственное бюджетное учреждение
«Сибирское отделение Российской академии наук».
630090, Россия, Новосибирск,
проспект Академика Лаврентьева, дом 17.

Издатель:
Федеральное государственное бюджетное учреждение
«Сибирское отделение Российской академии наук».
630090, Россия, Новосибирск,
проспект Академика Лаврентьева, дом 17.

Главный редактор:
академик РАН Валентин Николаевич Пармон.

Редакционный совет:
Академики РАН С.В. Алексеенко, И.В. Бычков, М.И. Воевода, А.П. Деревянко, Н.А. Колчанов, Д.М. Маркович, В.И. Молодин, В.М. Фомин, д.ф.- м.н. С.Р. Сверчков, генеральный директор ООО «Газпром трансгаз Томск» В.И. Бородин, генеральный директор АО «Академпарк» Д.Б. Верховод, заместитель полномочного представителя Президента РФ в СФО В.М. Головко, д.т.н., президент НГТУ НЭТИ, председатель Совета ректоров СФО Н.В. Пустовой, основатель АФК «Система» В.П. Евтушенков .

Редакционная группа:
Заместитель главного редактора Сергей Сверчков, Лариса Деева, Владимир Ларин, Андрей Соболевский, Татьяна Урбах, Любовь Батраева, Юлия Андреева.

Фото
авторов представленных материалов и из открытых источников.

Дизайн:
ФГБОУ ВО «Новосибирский государственный университет архитектуры, дизайна и искусств имени А.Д. Крячкова», ректор Багрова Наталья, арт-директор Чешева Татьяна, дизайнеры: Теряева Анна, Алиса Переверзева.

Свидетельство о регистрации СМИ ПИ № ФС 77–82311 от 03.12. 2021 г. выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций (Роскомнадзор).

Адрес редакции: 630090, Россия, Новосибирск,
проспект Лаврентьева, 17, тел.: (383) 217-45-78,
e-mail: l.batraeva@sb-ras.ru

Отпечатано в ООО «Процвет»
тел.: +7-951-854-86-18, +7-980-554-98-10,
e-mail: Vladimir-printvrn@mail.ru

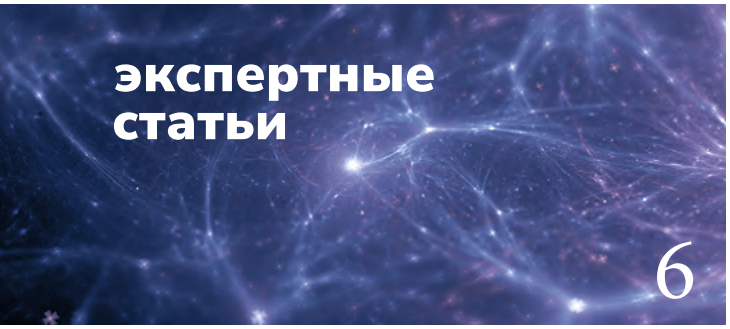
Подписано в печать 25.11.2023
Бумага мелованная. Печать офсетная.
Тираж 800 экз. Свободная цена.

Перепечатка материалов только с письменного разрешения редакции.
Изданию присвоен номер ISSN: 2782-4969

© Сибирское отделение РАН, 2023

в номере

стр. 4 Обращение главного редактора академика В. Н. Пармона



стр. 6 IT: эффекты консолидации

стр. 9 Сибирская IT-индустрия и искусственный интеллект

стр. 16 Франшиза бакалавриата компьютерных наук в Новосибирском государственном университете

стр. 24 Программный комплекс USPARS для решения систем линейных алгебраических уравнений в задачах моделирования

стр. 29 Unipro UGENE — свободный продукт для визуализации, анализа и обработки данных в молекулярной биологии и биоинформатике

стр. 34 Программный комплекс для разработки интеллектуальных систем поддержки принятия решений в абоформализованных областях

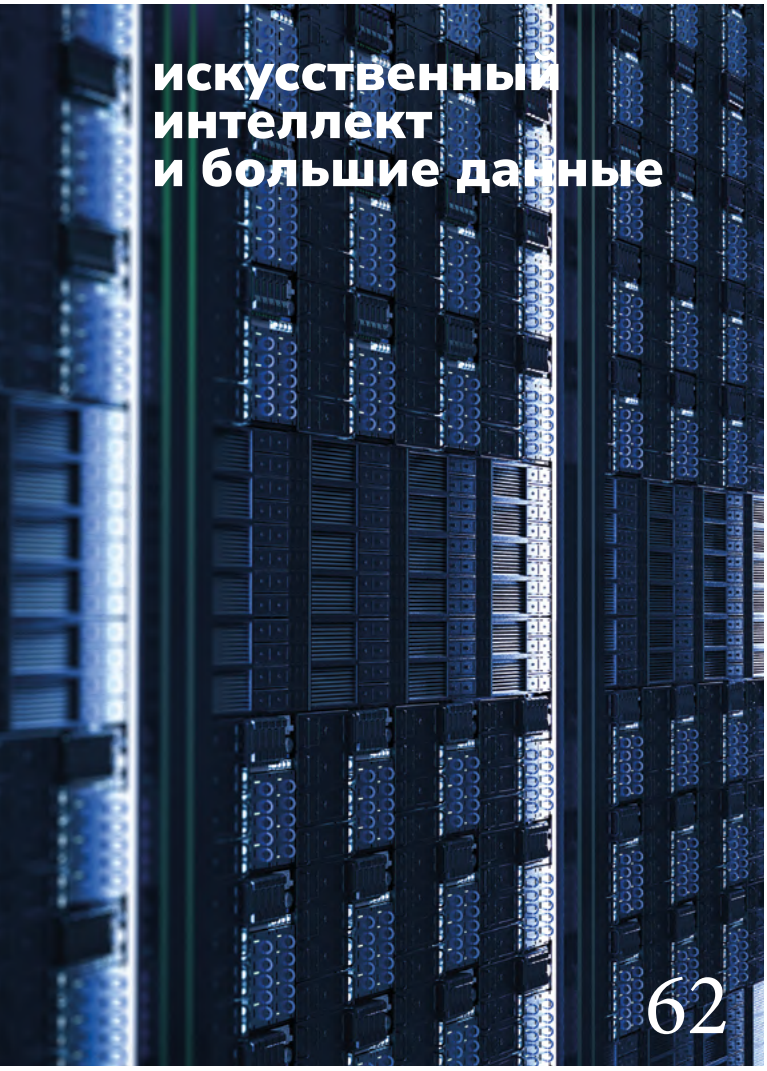
стр. 39 Технология прогноза урожайности картофеля с использованием дистанционного зондирования

стр. 41 Моделирование распространения эпидемий с учетом экономической и социальной ситуации в регионе

стр. 46 Клеточно-автоматная модель динамических процессов в газоугольном растворе

стр. 52 Оптические системы панорамной диагностики потоков «ПОЛИС»

стр. 56 Оптико-информационная диагностика для энергетики, промышленности и транспорта



стр. 64 Цифровая библиотека буддийского наследия: методы машинного обучения в изучении древних книжных памятников

стр. 70 Набор данных EMNIST TCyrillic для тренировки моделей машинного обучения распознавания символов тюркских письменностей, построенных на основе расширенной кириллицы

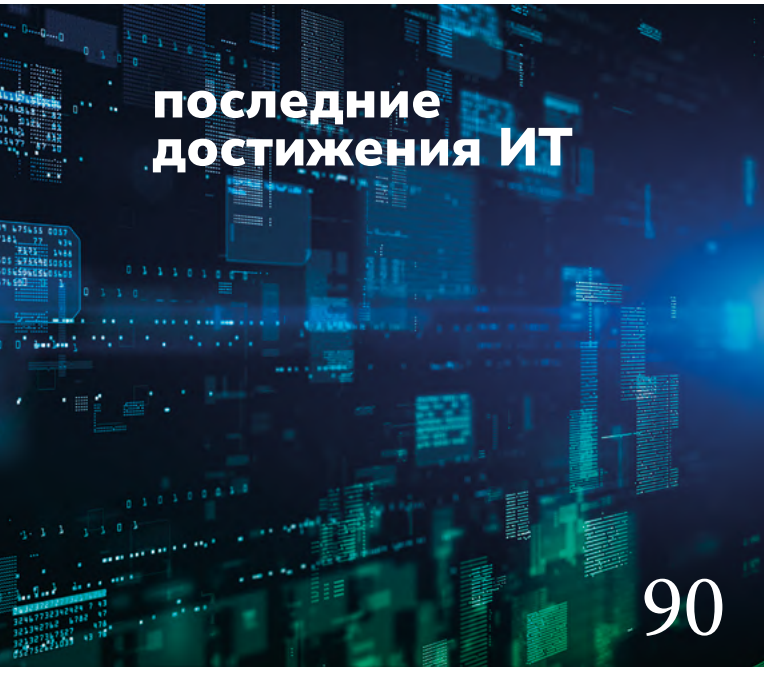
стр. 73 Нейросетевая система диагностики острого инсульта на основе компьютерного анализа КТ-изображений головного мозга

стр. 75 Интеллектуальная технология прогнозирования инсоляции

стр. 78 Технология для персонализированной диагностики нейроонкологических заболеваний с помощью искусственного интеллекта

стр. 82 Метод оценки эффективности применения геофизических исследований для решения геологоразведочных задач, основанный на применении искусственного интеллекта и построении трехмерной геолого-геофизической модели

стр. 85 Проблемы обеспечения правовой защиты персональных данных при использовании технологии искусственного интеллекта в медицинской деятельности



стр. 92 Staffcop Enterprise

стр. 96 Семейство программно-аппаратных комплексов обработки больших потоков данных в режиме реального времени

стр. 100 Разработка и верификация автоматных программ для систем управления

стр. 107 Методика оценки состояния и выбора структуры высоконадежной распределительной сети на основе применения меры неопределенности информации

стр. 111 FIS Platform: low-code платформа для автоматизации финансового сектора

Дорогие друзья!

У вас перед глазами — номер журнала, который может быть прочитан без вашего участия. Он посвящен информационным технологиям, достигшим в наше время невероятных высот, и так называемое компьютерное чтение стало рутиной. Искусственный интеллект сегодня способен реализовать прогнозы фантастов минувших десятилетий: рисовать картины, писать стихи и диссертации, управлять сложными технологическими процессами. Однако понятие информационных систем шире «нечеловеческого разума» — в конце концов, бумага и карандаш тоже являются информационной системой. Если же не утрировать, то IT — это целая Вселенная цифровых решений для всех, сегодня уже без исключений, сфер и отраслей человеческой деятельности.

Тотальность охвата наших практик IT-решениями подтверждает оглавление этого выпуска «Науки и технологии Сибири». Представленные здесь разработки относятся к клинической медицине и эпидемиологии, геофизике, процессам управления, кибербезопасности, сельскому хозяйству, энергетике, горному делу... И к научным исследованиям тоже: цифровые помощники ученых облегчают и ускоряют операции по поиску, систематизации и освоению информации. А в ряде ситуаций — делают за исследователя предварительную работу, что показано на примере перевода и расшифровки с помощью искусственного интеллекта древнетибетских рукописей.

IT можно назвать математикой, воплощенной в программировании. Это убедительно доказывает экспертная статья руководителя Международного математического центра в новосибирском Академгородке члена-корреспондента РАН Сергея Игоревича Кабанихина. Уже не удивляющие нас нейросети «учатся» за счет математики высочайшего уровня сложности, без которой каждый отдельно взятый искусственный нейрон не способен взаимодействовать с другими, а вся сеть — самосовершенствоваться и функционировать.

С нейросетью можно сравнить и само сибирское профессиональное IT-сообщество. Оно зародилось в процессе развития компьютерных компаний в 1990-х, когда в отдельных случаях стала возникать необходимость разработки оригинального софта. Первоначальный корпус специалистов для этой деятельности имелся благодаря наличию научных школ, выросших в непростое время гонений на кибернетику. В новосибирском Академгородке, прежде всего, работала яркая плеяда ученых: академики Гурий Иванович Марчук, Андрей Петрович Ершов, Михаил Михайлович Лаврентьев-старший, Николай Николаевич Яненко, Анатолий Семенович Алексеев, недавно скончавшийся Сергей Константинович Годунов, член-корреспондент АН СССР Геннадий Алексеевич Михайлов другие. Мировое признание заслужили их пионерские результаты в вычислительной и прикладной математике, кибернетике, в программировании и информатике, в математической геофизике и зарождавшихся компьютерных технологиях (термин появится позже).

Созданные этими корифеями научные школы отобразились в содержании учебных программ в НГУ и других вузах. На этой почве выросло первое поколение специалистов IT-отрасли (тоже раньше самого определения). В статье президента ассоциации «СибАкадемСофт» Ирины Аманжоловны Травиной описан процесс взаимодействия (включая неформальные) программистских «нейронов», предсказуемо связавшихся в «сети» — упомянутую выше ассоциацию и форум, одноименный названию этого выпуска журнала.

Зависимость прогресса IT-отрасли от человеческого фактора аксиоматична. Какие специалисты в нее придут — такими будут ее продукты и решения. Декан факультета информационных технологий Новосибирского государственного университета доктор физико-математических наук Михаил Михайлович Лаврентьев вместе с соавторами рассказывает об уникальной образовательной программе, совмещающей базовое и специализированное направления. Речь идет не о закрытом ноу-хау НГУ, а о продукте, предлагаемом для тиражирования. Примечательно, что эта программа разработана в НГУ совместно с коллегами из Хартфордширского университета (Великобритания), и даже в пору жесткой глобальной конфронтации проект остается международным.

Искренне желаю читателям этого номера успеха во всех начинаниях, профессиональных достижений, мира, здоровья и благополучия!

ВЕРНУТЬСЯ К СОДЕРЖАНИЮ

С уважением,
академик РАН Валентин Пармон

главный редактор издания
«Наука и технологии Сибири»,
председатель Сибирского отделения РАН,
вице-президент РАН

ВЕРНУТЬСЯ К СОДЕРЖАНИЮ

IT: ЭФФЕКТЫ КОНСОЛИДАЦИИ

Травина Ирина Аманжоловна

президент Ассоциации «СибАкадемСофт»

Отрасль информационных технологий является достаточно молодой. Но это и понятно, ведь сами ИТ-технологии начали бурно развиваться лишь во второй половине XX века.

Российским компаниям в нашей отрасли вообще не может быть больше 32 лет, поскольку только после распада СССР и образования РФ стало возможным свободно создавать коммерческие предприятия (хотя зачатком послужили кооперативы периода горбачёвской перестройки). Конечно, возраст ИТ-отрасли в России гораздо старше, так как существовали и продолжают существовать поныне множество отраслевых и научных организаций, а часть университетов в России обучала информационным технологиям и в советское время.

Появление ассоциаций в нашей отрасли началось не в 1991 году, а гораздо позднее, с тех пор, как ИТ-компании осознали, что только консолидировавшись они смогут донести свои задачи и проблемы до государства. Наиболее известными и долгоживущими ассоциациями в нашей отрасли являются АП КИТ (с 2002 года) и РУССОФТ (с 2004 года). Интересна и уникальна история создания нашей ассоциации «СибАкадемСофт».

Место рождения — новосибирский Академгородок. Почти за 10 лет после возникновения России как отдельного государства здесь возникло множество ИТ-компаний. Этому есть несколько объяснений. Главным фактором было то, что Новосибирский государственный университет, благодаря основателю сибирской школы информатики академику Андрею Петровичу Ершову, давал основы программирования на всех естественнонаучных факультетах: математикам, физикам, химикам, биологам, геологам и даже экономистам.

Выросли квалифицированные специалисты, обладающие знаниями в предметных научных областях и компетенциями в области программирования и алгоритмики. Поскольку спрос на ИТ-продукты рос пропорционально распространению компьютеров, а затем и других устройств, то почти любая компания могла начать работать и довольно быстро найти свою нишу на рынке. Вследствие этого в Академгородке появилось довольно много продуктовых ИТ-компаний, а также появились и команды, предлагающие услуги программирования любых задач.

Но Академгородок на стыке веков — это ещё полностью научный городок с 30 академи-

ческими институтами. Офисных центров не существовало, все компании так или иначе базировались в арендных помещениях институтов СО РАН. Айтишники уже тогда собирались, обсуждали в основном свои профессиональные темы. Были очень востребованы такие формы, к примеру, как пивной клуб. Но до серьезных задач не доходило, так как градус (в прямом и переносном смысле) обсуждения к концу заседания уже был не конструктивен. Помог случай. США довольно целенаправленно вели работу с различными компаниями из стран бывшего СССР. В рамках программы борьбы за демократию Министерство торговли США создало стажировку SABIT. Обычный формат группы — это 18 человек, топы компаний из одной отрасли, но из разных городов России и бывших советских республик. Однако в 2000 году была сформирована уникальная группа: руководители ИТ-компаний из Новосибирска, а точнее из Академгородка, то есть с одной территории. Стажировка длилась месяц. И участники довольно сильно подружились между собой. Посмотрели, как работают ИТ-компании в США, изучили работу бизнес-сообществ. И после этой поездки возникла идея создания своей ассоциации.

Поскольку мы все были родом из институтов и связаны с ними и с самим Сибирским отделением РАН, то идея была поддержана председателем СО РАН академиком Николаем Леонтьевичем Добрецовым. Более того, он помог убедить Администрацию Новосибирской области (в лице губернатора) и ректора НГУ стать соучредителями некоммерческого партнерства «СибАкадемСофт». Некоторое время ушло на согласование Устава. И вот 19 июня 2001 года была создана наша ассоциация. Соучредителями также выступили семь ИТ-компаний и исполнительная дирекция технопарка «Новосибирск». По сути — готовая кластерная структура. Правда, на тот момент подобные термины еще были не в ходу. Таким образом создалась одна из старейших российских айтишных ассоциаций, причем основной костяк участников — это компании-разработчики.

Уже в 2002 году мы решили провести собственный форум, который решал бы задачи, стоящие перед наши сообществом:

- лоббирование наших интересов на всех уровнях;
- продвижение нашей продукции;
- выявление узких мест и барьеров развития;
- кадровые вопросы.

Форум получил название «Сибирская индустрия информационных систем», коротко СИИС. Он прошел с успехом, на нём были сформулированы два основных направления, которые мы считали для себя наиболее важными. Это создание ИТ-парка и ИТ-деревни. Первое — потому что компаниям нужно было жизненное пространство: находиться на задворках и в подвалах институтов было уже неинтересно, требовались офисы другого уровня комфорта и обустройства. Второе было связано с тем, что Академгородок и тогда, и сейчас практически недоступен как место проживания молодежи. А кроме того, из-за существенных ограничений по строительству внутри Академгородка не хватало и более дорогого, но комфортного жилья для состоявшихся специалистов.

Следуя нашим стратегическим целям, в 2003–2005 годах «СибАкадемСофт» совместно с Экспертным советом по вопросам развития информационно-коммуникационных технологий при Полномочном представителе Президента России в СФО инициировали на федеральном уровне проект создания в Новосибирске технопарка. Проект был поддержан и губернатором Виктором Александровичем Толоконским, и главой Минкомсвязи РФ Леонидом Дододжоновичем Рейманом. Сегодня можно констатировать, что Академпарк во многом состоялся из-за проактивной позиции «СибАкадемСофта». И сейчас ассоциация тесно сотрудничает с Академпарком по вопросам развития ИТ-отрасли, помогает в инкубировании ИТ-стартапов, совместно проводит различные мероприятия.

А форум СИИС стал традиционным. В текущем 2023 году он был проведён в 13-й раз, на своей родной площадке в Академпарке. 17 мероприятий, более 600 участников, важные итоги в виде возрождения ИТ-кластера в Новосибирской области и старта проекта «СмартСити». И о том и о другом чуть ниже.

В копилке «СибАкадемСофта» есть ещё одно крупное достижение. Подготовленная нашей ассоциацией заявка на программу кластерного развития заняла одно из первых мест в федеральном конкурсе Минэкономразвития РФ и послужила основой Инновационного объе-

диненного кластера информационных и биофармацевтических технологий Новосибирска, который решением Правительства РФ по результатам конкурса в августе 2012 года включен в число 14 приоритетных. Реализация кластерных проектов позволила привлечь из федерального бюджета около 700 млн рублей субсидий на цели их создания и становления.

На данный момент основными задачами нашей ассоциации являются:

- налаживание кластерного взаимодействия участников;
- консолидированное продвижение товаров и услуг компаний-участников «СибАкадемСофта» на российский и зарубежный рынки;
- влияние на внутреннее регулирование ИТ-рынка;
- взаимодействие с органами власти любого уровня в целях отстаивания интересов сибирских разработчиков программных продуктов и содействия развития малого и среднего бизнеса в Сибирском макрорегионе;
- кадровая и методическая поддержка образовательных учреждений, готовящих специалистов в сфере ИТ;
- решение жилищных проблем сотрудников ИТ-компаний и обеспечение комфортных условий труда, чтобы по возможности сдерживать отток специалистов за рубеж.

Исходя из этих задач, одним из приоритетных проектов для ассоциации на данный момент стал проект «СмартСити» в Академгородке. Он не просто призван решить оставшиеся по наследству проблемы — это попытка создать инновационно-научный центр мирового уровня. Так уж сложилось, что наличие уникального научного кластера в виде более 30 научных организаций и очень сильного университета превратило Академгородок в дом для сотен инновационных компаний. Спрос на кадры растет, но их не хватает, так как большая часть выпускников оказывается зачастую на Западе (России и просто зарубежном Западе). Перспективные абитуриенты Академгородка уезжают учиться в Москву и редко возвращаются. Проект «Смарт-Сити» — ответ на эти вызовы, и судя по отзывам школьников, которым мы его презентовали, это то, что им нужно. Они ищут образ будущего, с которым им интересно связать свою жизнь, и в наших силах попытаться этот образ воплотить в реальности.

В текущем номере журнала «Наука и технологии Сибири» представлены продукты и технологии наших компаний. Хочу сказать, что многие из них успешно конкурируют на мировом рынке. Причем до февраля 2022 года им проще было работать с зарубежными покупателями, чем с российскими. Меньше барьеров, проще тендерные процедуры, удобные и раскрученные маркетплейсы. Но в государстве произошел кардинальный поворот к наращиванию собственного пула технологий, независимого от иностранных, зачастую недружественных, вендоров. Это дало возможность многим компаниям-участникам ассоциации нарастить объем сделок внутри России. Ведь мы всегда были чисто отечественными производителями и существовали, как правило, не благодаря, а вопреки всему. Компаниям удалось сохранить у себя исключительные компетенции, которые могут помочь решить государственную задачу импортонезависимости. Не секрет, что многие зарубежные продукты, к которым привыкли миллионы пользователей в России, разрабатывались десятки лет. Чтобы их заменить, нужны достаточно крупные вложения, которые не под силу самим разработчикам. Поэтому будем интенсивно продвигаться в направлении государственной поддержки этой деятельности.

Я с оптимизмом смотрю с будущее. Возможно, это связано с тем, что я всю свою жизнь работаю в развивающейся отрасли. И постоянно общаюсь с молодыми компаниями, которые появляются у нас в технопарке, вижу, как они растут и прогрессируют.

Оптимизма добавляет и то, что исследовательские организации в Академгородке не просто сохранили свои компетенции — многие из них очень востребованы не только в России, но и вне её. Меня окружает интеллектуальная элита страны, которая вдали от высоких начальственных кругов по-прежнему имеет возможность реализовывать свой интеллектуальный и культурный потенциал. И номер журнала, который вы взяли в руки, — еще одно подтверждение этому.

СИБИРСКАЯ IT-ИНДУСТРИЯ И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

Эксперты Высшей школы экономики на основе данных Росстата впервые оценили вклад российской информационной индустрии в экономику страны. По данным за 2016 год, доля информационной индустрии в России составила 3,3% ВВП.

Кабанихин Сергей Игоревич

доктор физико-математических наук, член-корреспондент РАН, директор Международного математического центра ИМ СО РАН

1. Информационная индустрия

В большинстве развитых стран этот показатель находится на уровне 5%, в Великобритании, Венгрии, Ирландии, Республике Корея, США, Японии достигает 7%.

Сибирская IT-индустрия занимается разработкой, производством и внедрением информационных систем в Сибирском макрорегионе России. Она включает в себя компании, специализирующиеся на создании программного обеспечения, разработке баз данных, информационной безопасности, интернет-технологиях. Центральный по местоположению и один из крупнейших по территории субъектов РФ имеет исключительные возможности для развития инфраструктуры и организации новых технологических центров, а богатые природные ресурсы (древесина, нефть, газ) могут быть использованы для инновационных решений в области информационных систем. Поскольку в Сибири активно работают технопарки, инкубаторы и акселераторы, которые поддерживают развитие информационных технологий и способствуют созданию новых проектов, IT-индустрия имеет большой потенциал для развития и привлечения инвестиций, может и должна стать одной из ключевых отраслей экономики региона, способствуя ее модернизации.

Отметим ряд компаний, активно работающих по различным направлениям IT-индустрии:

- «Ланит-Терком» — разработка и внедрение информационных систем для государственного сектора, банков, телекоммуникаций и других отраслей.
- «Сибирские решения» — программное обеспечение и информационные системы для

автомобильной промышленности, энергетики и транспорта.

- «АйТеко» — разработка и внедрение информационных систем для банков и финансовых учреждений.

- «Сибинформсистемы» — услуги в области информационной безопасности, разработки программного обеспечения и консалтинга.

Отметим также ЦФТИ (Центр физических и технических исследований) в новосибирском Академгородке, в котором проводятся фундаментальные и прикладные исследования в области физики, математики, информатики, биологии и других научных дисциплин. ЦФТИ насчитывает более 1300 научных сотрудников и является также центром образования и подготовки специалистов в различных областях науки и техники, предлагая образовательные программы для студентов, аспирантов и молодых ученых.

Очень важна поддержка IT-индустрии со стороны региональных инкубаторов и технопарков, таких как Академпарк в Новосибирске и «Томское поколение» в Томске.

Среди основных задач, которые призвана решать IT-индустрия, отметим следующие:

1. Автоматизация бизнес-процессов.

2. IT-индустрия помогает собирать, хранить и обрабатывать данные, что позволяет принимать обоснованные решения на основе анализа информации.

3. IT-индустрия облегчает коммуникацию и обмен информацией между сотрудниками, отделами и партнерами организации.

4. IT-индустрия позволяет организациям защищать свои данные от несанкционированного доступа и утечек.

5. IT-индустрия дает организациям возможность улучшить взаимодействие с клиентами, предоставляя им удобные способы заказа товаров или услуг, обратной связи и поддержки, помогает организациям эффективно управлять своими ресурсами, такими как финансы, материалы и персонал.

6. IT-индустрия позволяет организациям анализировать данные и проводить прогнозирование, что помогает оптимизировать бизнес-процессы, прогнозировать спрос и предложение, а также управлять рисками.

В заключение отметим возрастающую роль технологии искусственного интеллекта (ИИ) при анализе потребительского рынка (маркетинг) в случае, когда классические методы прогнозирования отклика потребителей могут быть недостаточно точны и могут быть дополнены прогнозирующей нейросетевой системой. Исследование спроса на основе таких систем незаменимо для компании в условиях конкуренции, поскольку позволяет поддерживать постоянный контакт с потребителями через обратную связь. В настоящее время корпорации проводят опросы потребителей, позволяющие выяснить, какие факторы являются для них решающими при покупке данного товара или услуги, почему в некоторых случаях предпочтение отдаётся конкурентам и какие товары потребитель хотел бы увидеть в будущем. Анализ результатов этих опросов при помощи нейросетевой системы позволяет выявить сложные зависимости между факторами спроса, прогнозировать поведение потребителей при изменении маркетинговой политики, находить наиболее значимые факторы и оптимальные стратегии рекламы и, таким образом, искать наиболее перспективный класс потребителей какого-либо товара или услуги.

2. Технологии искусственного интеллекта – основа IT-индустрии

Весь набор технологий и методов ИИ позволяет компьютерным системам имитировать и выполнять задачи, требующие значительных человеческих затрат. ИИ активно используется в информационных системах для автоматизации процессов, анализа данных, принятия решений и предоставления персонализированных услуг, в том числе для обработки больших объемов данных и выявления скрытых закономерностей и тенденций, в автоматической классификации и категоризации информации. ИИ помогает в создании интеллектуальных агентов, которые могут выполнять определенные задачи автономно, что позволяет сотрудникам организации сосредоточиться на более сложных задачах, в то время как ИИ выполняет рутинные и повторяющиеся задачи. Один из примеров — система управления клиентскими отношениями (CRM), которая использует алгоритмы машинного обучения и анализа данных для предсказания потребностей клиентов и предоставления

персонализированных рекомендаций и предложений. Перспективно также использовать методы ИИ для прогнозирования спроса на продукцию, оптимизации расписания производства и управления запасами.

3. Машинное обучение, нейронные сети и математическое моделирование

Машинное обучение — одно из важнейших направлений создания и развития систем ИИ, а нейронные сети позволяют решать задачи машинного обучения практически во всех областях современной прикладной математики, включая распознавание образов, оптимизацию и экстраполяцию функций, многомерных уравнений в частных производных, регуляризацию обратных и некорректных задач математической физики.

Обучение линейных нейронных сетей осуществляется путем корректировки весов нейронов матрицы весов с целью минимизировать ошибку между прогнозируемыми и фактическими значениями в обучающих данных. Для этой цели используются методы оптимизации (в основном градиентный спуск). Перед обучением необходимо определить функцию потерь, которая измеряет разницу между прогнозируемым значением и фактическим значением. Обычно используется стандартная ошибка или средняя абсолютная погрешность. После определения функции потерь можно приступать к обучению. На каждой итерации градиентного спуска вес нейрона обновляется в направлении, противоположном градиенту функции потерь. Это позволяет свести к минимуму ошибки и повысить предсказательную способность сети. Обучение линейных нейронных сетей может быть довольно быстрым и эффективным, особенно для регрессионных задач с большим количеством функций.

В настоящее время наиболее развитыми классами нейросетей (NN) являются сверточные NN (Convolutional Neural Networks, CNN) — классификация изображений, распознавание объектов и детектирование лиц; рекуррентные НС (Recurrent Neural Networks, RNN) — машинный перевод, распознавание речи и анализ текстов; глубокие NN (Deep Neural Networks, DNN) — распознавание рукописных цифр и прогнозирование финансовых показателей; генеративно-состязательные NN (Generative Adversarial Networks, GAN) — создание изображений, музыки и видео; автоэнкодеры (Autoencoders) — сжатие данных и извлечение признаков, уменьшение размерности данных и реконструкция изображений.

Физически обоснованные NN (Physics-Informed Neural Networks, PINN) и основанные на данных NN (Data Driven NN) — наиболее эффективны для задач моделирования и анализа физических процессов (расчет напряжений и деформаций в материалах, прогнозирование погоды и анализ экономических данных). Они сочетают в себе преимущества физических моделей и машинного обучения для получения более точных и универсальных результатов. При этом PINN и Data Driven NN могут использоваться для моделирования в различных областях, включая информационные системы, например, для прогнозирования поведения пользователей на сайте или для оптимизации процессов обработки данных.

PINN и Data Driven NN могут использоваться как взаимосвязанные сети для решения сложных задач моделирования. Например, PINN моделирует физический процесс, а Data Driven NN анализирует входящие в этот процесс данные. Комбинируя эти сети, можно получить более точные и эффективные модели (биология, науки о Земле, социология, эпидемиология, экология, экономика — примеры такого подхода, развиваемого в Международном математическом центре Академгородка).

В этой связи одна из важных задач для математиков — обоснование результатов работы алгоритмов машинного обучения.

Мы рассмотрим простейший класс НС — линейные. Линейные нейронные сети (LNN) — это класс нейронных сетей, в которых каждый нейрон имеет только один выходной сигнал и выполняет линейные операции. В линейной нейронной сети каждый нейрон получает входные данные, умножает их на соответствующий вес и суммирует результаты. Затем результат передается на выход нейрона [6,7].

Приведем простейший пример линейной нейросети.

Пусть

$$h = \sum_{i=1}^{N} g_i q_i + b\,,\ i = 1,2,\cdots N,\ y = f(h)\,,$$

где h — выходной вектор, g — входной вектор, q — вектор весов, b — параметр смещения, f — функция активации.

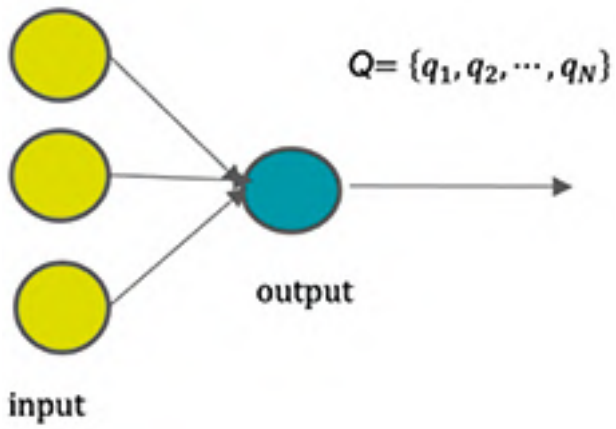


Рис. 1. Структура линейной нейронной сети.

В более сложных случаях векторы h, g и q заменяются матрицами H, G и Q.

Рассмотрим для примера линейную задачу машинного обучения (ЛЗМО), в которой требуется определить матрицу Q параметров нейронной сети по обучающим выборкам G и H.

Покажем, как ЛЗМО может быть сведена к системе линейных алгебраических уравнений (СЛАУ).

Даны два множества векторов:

$$G = \{g^{(1)}, g^{(2)}, \cdots, g^{(k)}\}, H = \{h^{(1)}, h^{(2)}, \cdots, h^{(k)}\},$$

и матрица Q, такие, что

$$Q g^{(k)} = h^{(k)},\ k=1,2,\cdots,K. \qquad (1)$$

Требуется восстановить матрицу Q. В общем случае матрица Q прямоугольная $Q= Q_{MN} = \{q_{mn}\}$. m — номер строки. В этом случае $g^{(k)} \in R^N$, $h^{(k)} \in R^M$.

Пусть G_{MN} — матрица, составленная из столбцов $g^{(1)}, g^{(2)}, \ldots, g^{(k)}$, а H^{MN} — из столбцов $h^{(1)}, h^{(2)}, \ldots, h^{(k)}$. Тогда относительно Q_{MN} получим матричную СЛАУ.

$$Q_{MN} \cdot G_{NK} = H_{MK}. \qquad (2)$$

Если M = N = K, а матрица G_{NK} хорошо обусловлена, то ЛЗМО решается просто

$$Q_{KK} = H_{KK} G_{KK}^{-1}. \qquad (3)$$

а главный вывод: множество обучающих векторов G_{KK} должно быть линейно независимым.

Для изучения оставшихся случаев отметим, что из (2) несложно получить следствие

$$G_{NK}^{T} \cdot Q_{MN}^{T} = H_{MK}^{T}. \qquad (4)$$

Обозначая $G_{NK}^{T} = A_{KN}$, $H_{MK}^{T} = F_{KM}$, перепишем (4) в виде

$$A_{KN} \cdot Q_{MN}^{T} = F_{KM}. \qquad (5)$$

Исследуем сначала отдельно СЛАУ для каждого столбца q^n, матрицы Q_{MN}^{T}.

$$A_{KN} \cdot q_N^n = f_K^n. \qquad (6)$$

Очевидно, что вопросы разрешимости всех этих СЛАУ, n=1,2,⋯, N, зависят в основном от матрицы A_{KN}. Рассмотрим систему

$$A_{KN} \cdot q_N^n = f_K^n. \qquad (7)$$

По теореме о сингулярном разложении

$$A_{KN} = U_{KK}\, \Sigma_{KN}\, V^*_{NN}$$

Здесь U_{KK} и V^*_{NN} ортогональны, а Σ_{KN} =diag $\{\sigma_1, \sigma_2, \cdots, \sigma_\rho, 0, \cdots, 0\}$, $\rho = \min\{K,N\}$

Следовательно, из (7) вытекает

$$\Sigma_{KN}\, V^*_{NN}\, q_N^n = U^*_{KK} f_K^n. \qquad (8)$$

Величина σ_ρ определяет степень устойчивости определения (комбинаций) параметров (матрицы Q) вектора q_N^n. Если $\sigma_\rho > 0$, то

$$(V^*_{NN} q_N^n)_j = 1/\sigma_j (U^*_{KK} f_K^n)_j,\ j=1, 2, \cdots, \rho. \qquad (9)$$

Лемма. При положительном σ_ρ ЛЗМО позволяет определить однозначно линейные ρ комбинации (9) матрицы Q.

Таким образом, в общем случае ЛЗМО не может быть решена однозначно и тем более устойчиво. Улучшить качество построенной нейросети можно, увеличивая обусловленность матрицы G_{NK}.

Более общая картина вариантов ЛЗМО после сведения к СЛАУ отражена в диаграмме [5].

ДИАГРАММА

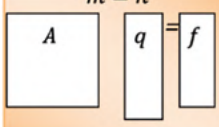
(СЛАУ) $Aq = f$, $A - m{\times}n$-матрица $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$.

$f \in \mathbb{R}^m, q \in \mathbb{R}^n$; $R(A) = \{f \in \mathbb{R}^m\colon\ \exists q \in \mathbb{R}^n\colon\ Aq = f\}$; $Q_f(A) = \{q \in \mathbb{R}^n\colon\ Aq = f\}$.

$|Q_f(A)|$ – количество решений системы $Aq = f$ при заданных A и f

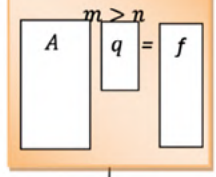
$\operatorname{rank} A = \min(m, n) := p$

$m = n$



$|Q_f(A)| = 1, \forall f \in \mathbb{R}^m$, т. е. *существует единственное* **точное решение** $q_{\text{т}}$ для

$m > n$



$R(A) \neq \mathbb{R}^m$
$|Q_f(A)| \leq 1, \forall f \in \mathbb{R}^m$, т. е. если **точное решение** существует, то оно *единственно*

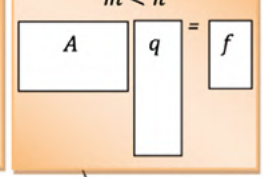
$f \in R(A)$

$A^*Aq = A^*f$, где A^* - сопряженная к A матрица ($A^* = \overline{A^{\text{T}}}$; $A^* = A^{\text{T}}$, если $A \subset \mathbb{R}^n{\times}\mathbb{R}^n$)

$f \notin R(A)$

$|Q_f(A)| = 0$

$m < n$



$|Q_f(A)| = \infty, \forall f \in \mathbb{R}^m$, т. е. **точное решение** *существует* для любого $f \in \mathbb{R}^m$

существует единственное **нормальное решение** $q_{\text{н}} := \arg\min_{q \in Q_f(A)} \|q\|$

$\operatorname{rank} A < \min(m, n)$

$R(A) \neq \mathbb{R}^m$, $|Q_f(A)| \neq 1$

$f \in R(A)$

$|Q_f(A)| = \infty$

$f \notin R(A)$

$|Q_f(A)| = 0$

существует **псевдорешение** $q_{\text{п}} := \arg\min_{q \in \mathbb{R}^n} \|Aq - f\|$

$Q_f^{\text{п}}(A)$ – множество псевдорешений системы $Aq - f$, $|\ Q_f^{\text{п}}(A)\ | \geq 1$.

$|\ Q_f^{\text{п}}(A)\ | > 1$

существует единственное **нормальное псевдорешение** $q_{\text{нп}} := \arg\min_{q \in Q_f^{\text{п}}(A)} \|q\|$

$|Q_f(A)| = 1$

A (или A^*A) хорошо обусловлена, т. е. число обусловленности $\mu(A) := \sup_{\substack{x \neq 0 \\ \xi \neq 0}} \left\{ \frac{\|Ax\|\|\xi\|}{\|A\xi\|\|x\|} \right\}$ относительно

решение неустойчиво к возмущениям правой части f, т. е. задача определения точного решения системы $Aq = f$ корректна.

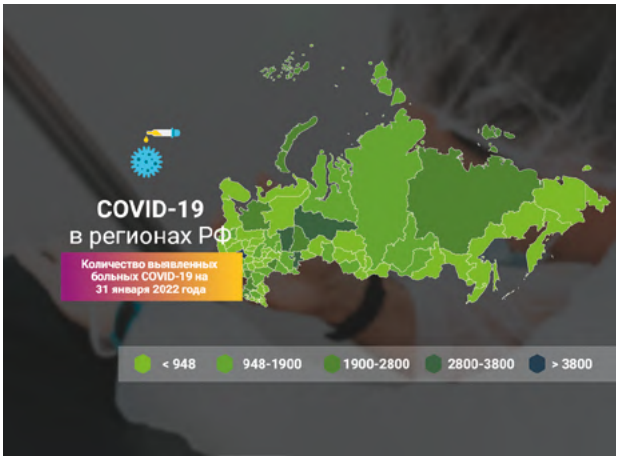
A (или A^*A) плохо обусловленная матрица, т. е. число обусловленности

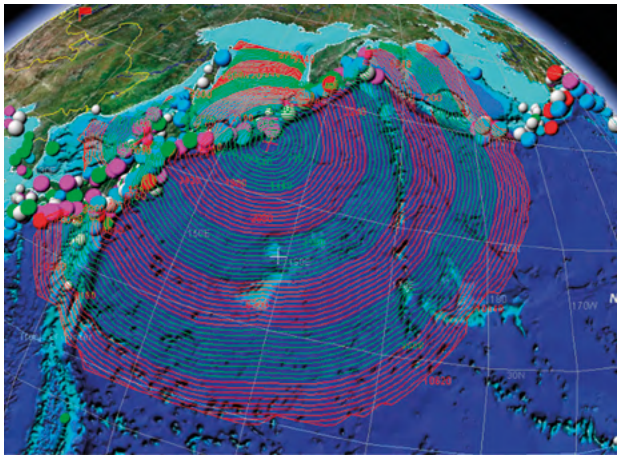
решение *неустойчиво* к возмущениям правой части f, т. е. задача нахождения точного решения системы $Aq = f$ **некорректна**.

4. Заключение

Технологии умных городов, практики применения искусственного интеллекта, разработки для госсектора, импортозамещение в ИТ-сфере — все эти новые возможности цифровой индустрии в условиях глобальных вызовов были темой обсуждения экспертов в Новосибирской области на XIII форуме «Сибирская индустрия информационных систем» (СИИС). Убежден, что основой существующих и новых технологий ИИ является математика. Именно поэтому одним из главных направлений деятельности Международного математического центра в Академгородке (консорциум на основе Института математики имени С. Л. Соболева СО РАН и Новосибирского государственного университета) является исследование применения и обоснования ИИ с помощью фундаментальных математических дисциплин. Скажу о работе только одной из более чем десяти научных групп матцентра «Обратные задачи естествознания». В буклете, который готовит замдиректора ММЦ Ольга Криворотько, кратко говорится об основных направлениях работы этой научной группы. И таких страничек будет более десяти.

Отметим, что больше половины сотрудников ММЦ — молодые ученые, аспиранты и студенты. Так что с уверенностью можно сказать, что молодые ученые Новосибирского научного центра СО РАН уже имеют хороший задел в сфере IT-индустрии и ИИ, и впереди у них большое будущее [1–4].





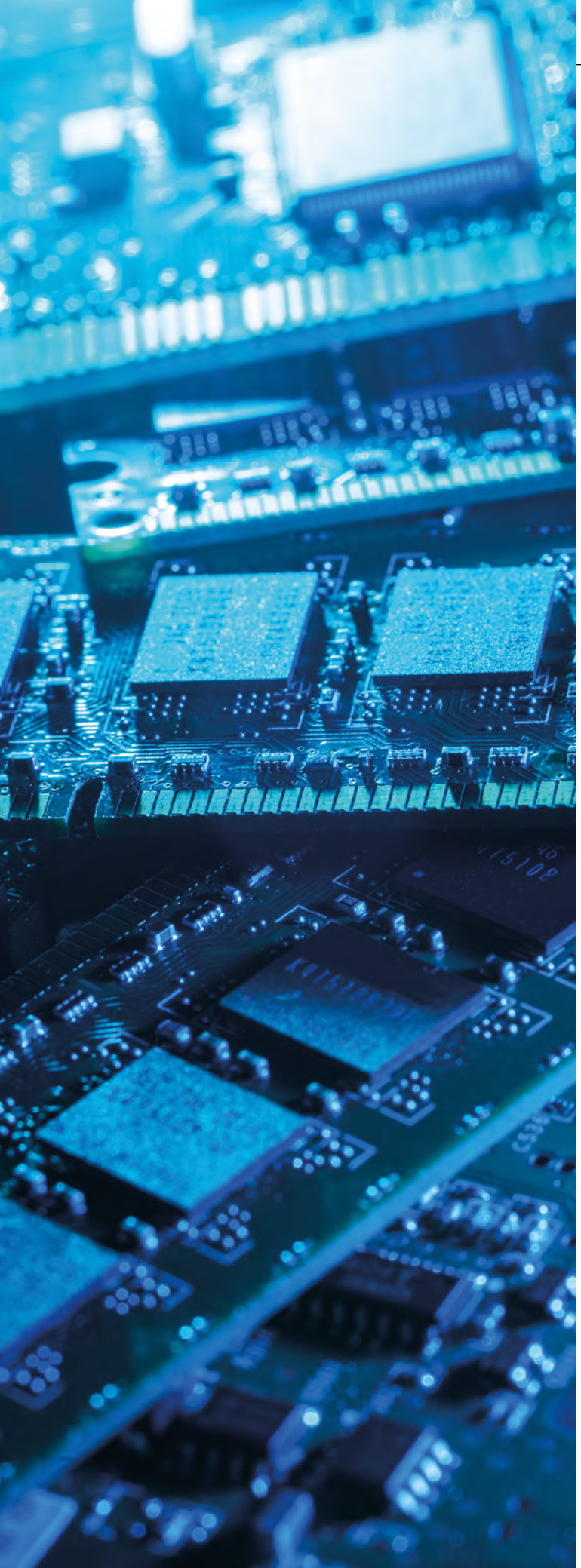
Организуемые конференции

Школа-конференция «Теория и численные методы решения обратных и некорректных задач», Новосибирск

Школа «Обратные задачи и машинное обучение», Сочи

Международная конференция «Современные проблемы обратных задач», Новосибирск

BGRS Symposium “Mathematics, bioinformatics and systems computational biology of COVID-19”, Novosibirsk



Обратные задачи естествознания

Основные направления исследований:

- Создание и анализ математических моделей взаимосвязанных экономических, эпидемиологических и экологических процессов с учётом социологической ситуации в регионах.
- Машинное обучение, Data-Driven, Physics Informed Neural Networks.
- Прямые и обратные задачи в науках о Земле, жизни и человеке.

Основные результаты исследований:

- Платформа для расчета сценариев развития эпидемий, социальных настроений.
- Программный комплекс моделирования природных и техногенных катастроф, включая цунами, вулканы, землетрясения.

Цитируемая литература

1. Тихонов А. Н. О методе регуляризации. Докл. АН СССР. 1963. Т. 151. № 3. С. 501–504.
2. Годунов С. К., Антонов А. Г., Кирилюк О. И., Костин В. И. Гарантированная точность решения систем линейных уравнений в евклидовых пространствах. Новосибирск: Наука. Сиб. отделение, 1992.
3. Кабанихин С. И. Обратные и некорректные задачи. Сибирское научное издательство, Новосибирск, 2009.
4. Галушкин А. И. Теория нейронных сетей: Учеб. пособие для вузов. М.: Радиотехника, 2000.
5. Кабанихин С. И. Обратные задачи и искусственный интеллект. Успехи кибернетики. 2021. Т. 2. № 3. С. 33–43.
6. Unsupervised Learning of Equivariant Structure from Sequences by Takeru Miyato, Masanori Koyama, Kenji Fukumizu, NeurIPS 2012.
7. Workshop on Functional Inference and Machine Intelligence (FIMI 2023).

Научные партнеры:

Российский научный фонд

Министерство науки и высшего образования Российской Федерации

Биотехнологическая компания BIOCAD

Huawei Technologies Co. Ltd.

ФРАНШИЗА БАКАЛАВРИАТА КОМПЬЮТЕРНЫХ НАУК
В НОВОСИБИРСКОМ ГОСУДАРСТВЕННОМ УНИВЕРСИТЕТЕ

Лаврентьев Михаил Михайлович

доктор физико-математических наук, профессор, декан ФИТ НГУ, заместитель директора по научной работе Института автоматики и электрометрии (ИАиЭ) СО РАН, руководитель Лаборатории программных систем машинной графики ИАиЭ СО РАН

Базовый уровень высшего образования в области компьютерных наук представлен бакалавриатом, традиционно составляющим четырехлетнюю программу.

В большинстве случаев программа бакалавриата естественно подразделяется на первую и вторую фазу. На первой фазе в течение двух лет обучение происходит без специализации и предметов по выбору. Вторая фаза должна обеспечить возможность специализации, большого количества предметов по выбору и выпускной квалификационной работы. Опыт показывает, что первая фаза практически не зависит от специфики второй: во всех компьютерных областях необходимо рабочее знание классической инженерной математики (анализ, линейная алгебра, геометрия, и т.д.) и специфической дискретной: основ логики, множеств, отношений, графов, а также комбинаторики и теории вероятности и статистики. Для всех областей компьютерных технологий необходимо уверенное владение методологией и инструментарием программирования в современном смысле: императивное и декларативное программирование, основы программного конструирования и организации разработки программного обеспечения (ПО). Наконец, современное состояние компьютерных наук и технологии немыслимо без глубокого понимания архитектуры и функционирования вычисляющих платформ. На втором году обучения возможны небольшие вариации материала для лучшей адаптации к специфике конкретных вариантов второй фазы. Естественно, что все эти вариации должны включаться в общую схему подачи материала второго года обучения.

Традиционная система подготовки ИТ-кадров классическими университетами России не вполне адекватна ожиданиям работодателей о получении квалифицированных выпускников вуза, готовых к работе без существенной переподготовки. Создание прочной теоретической базы неизбежно влечет за собой сравнительно позднее знакомство с профессиональными компонентами.

Необходимость изучения дисциплин, не связанных напрямую с ИТ-отраслью, демотивирует студентов младших курсов. Такой подход не соответствует лучшим мировым практикам. Кроме того, далеко не все университеты обладают достаточным количеством

кадров, способных преподавать необходимые теоретические дисциплины на мировом уровне. В большинстве технических (политехнических) университетов образование грешит обратным — тренингом студентов на освоение конкретных текущих инструментов и технологий, отдельных пакетов программного обеспечения (ПО). Это можно характеризовать как фиктивный прагматизм, оборачивающийся зависимостью от быстро устаревающих методов и практик. Выпускник настраивается на нетворческое, ремесленное отношение к профессии, отличающееся пренебрежением к глубоким пониманиям ключевых концепций и структур, отсутствием критического отношения к полученным знаниям, умениям и навыкам.

Совершенно естественно, что университеты РФ стремятся сохранить наработанные образовательные походы, выверенные образовательные программы, учебно-методические материалы и профессорско-преподавательский состав. Многие преподаватели годами ведут занятия по своим традиционным предметам и не видят необходимости что бы то ни было менять в учебном процессе. Любые изменения требуют дополнительной работы и затрат времени, часто весьма значительных. Как правило, вузы не обладают достаточными ресурсами, чтобы в должной мере мотивировать дополнительные усилия преподавателей. Изменения же, диктуемые «сверху», выполняются без энтузиазма, не творчески, а лишь чтобы формально отчитаться.

Плавное изменение сложившейся ситуации потребует многих лет, что противоречит необходимости решения насущной задачи обеспечения экономики РФ должным количеством квалифицированных ИТ-кадров в ближайшее время. Насильственное внедрение новых методов и подходов не представляется эффективным, поскольку «итальянские забастовки» в системе образования почти невозможно обнаружить и еще труднее преодолеть.

Инициативная группа, состоящая из преподавателей НГУ и Хартфордширского университета (Великобритания), разработала и внедрила на Факультете информационных технологий НГУ уникальную образовательную программу, представляющую обе фазы базовой программы бакалавриата по компьютерным наукам. По программе уже состоялись два выпуска бакалавров, и результаты (успешность защит, количество красных дипломов, конкурс на этот профиль и популярность среди абитуриентов) говорят сами за себя. Достаточно сказать, что по уровню подготовки в части как теории, так и практической проектной работы она успешно конкурирует с университетами категории ТОП 20 в области компьютерных наук (такими как Университет Эдинбурга и Имперский колледж в Великобритании), при этом оставаясь полностью поддержанной местными кадрами, учебными материалами и автоматизированными системами обучения и проверки знаний.

Эта программа отличается качественно лучшим уровнем согласования отдельных дисциплин в единую общую систему. Другое ее достижение — степень зрелости и «обкатки» учебных материалов (в особенности банков задач и упражнений). Как результат достигается повышенная мотивация студентов, а раскрытие их творческого потенциала находится на беспрецедентном уровне. В качестве примера можно привести регулярное инициирование студентами младших курсов собственных

Шафаренко Александр Владимирович

профессор, советник ректора НГУ, эксперт в области Computer Science и Software Engineering, информационной и кибербезопасности, руководитель нескольких крупных успешных проектов

Держо Марина Анатольевна

старший преподаватель ФИТ НГУ, специалист в области образовательной информатики, UX/UI

проектов с амбициозными целями, не связанных напрямую с изучаемыми материалами, а являющихся следствием желания решить необычную и интересную задачу. Отметим, что первая фаза образовательного продукта уже успешно используется для поддержки курсов другого подразделения НГУ — Института интеллектуальной робототехники, что демонстрирует универсальность ее базовых компонент.

Программа уникальна и не имеет аналогов ни в России, ни в Великобритании и других западных странах. В этом смысле деятельность носит коллаборативный характер, не сводящийся ни к импорту, ни к экспорту образовательных продуктов. Программа сочетает в себе лучшие стороны российской и западной систем преподавания.

Со стороны РФ это:

- глубокие курсы императивного программирования в сочетании с дискретной математикой, сопровождаемые решением большого количества задач: от легких до критической сложности. Используется автоматическая система проверки представляемых студентами решений (1 курс);
- курс декларативного программирования (язык Haskell) с элементами дискретной математики, позволяющий овладеть обеими парадигмами программирования (императивной и декларативной) и осознать в сравнении их преимущества и недостатки (1 курс);
- команда разработчиков с научным бэкграундом (преимущество новосибирского Академгородка) и опытом работы в успешных ИТ компаниях;
- междисциплинарность подхода.

Со стороны Великобритании:

- вводные предметы по компьютерной архитектуре, схемотехнике и низкоуровневому программированию, включая систему автоматической проверки разработанных схем и написанных студентами программ (1 курс);
- разработанные групповые проекты 1-го курса;
- координация создания программы и мониторинг ее апробации.

Основные инновационные характеристики этой образовательной программы состоят в следующем:

1. Широкое введение в теоретическую часть одновременно с прохождением профессиональных дисциплин вместо опережающего преподавания теории, как это принято в (пост)советских университетских программах. Академически состоятельные общие курсы ведутся параллельно профессиональным с горизонтальными связями («примеры из жизни») и понедельной координацией подачи материала между дисциплинами.

2. Несмотря на наличие именно введения в теорию, обеспечена высокая степень теоретической поддержки, осуществляемая на более старших курсах, что приводит к достижению отличного уровня компетентности выпускников в части методов, моделей и техники решения задач. По завершении обучения выпускники получают полное (неупрощенное) преподавание теории и общих академических дисциплин с высокой научной строгостью без скидок на практическую ориентацию курсов. Это достигается, в частности, за счет возвратов к дисциплине на новом уровне на старших курсах в требуемом объеме, когда это требуется для практических целей.

3. Преподавание профессиональных дисциплин с 1 курса с упором на решение задач и проектную деятельность (практико-ориентированный и проектный подходы). Опережающее преподавание профессиональных дисциплин до полного завершения общих курсов. Существенным является проектный компонент, то есть выполнение групповых и индивидуальных проектов каждый год, начиная непосредственно с первого курса, с целью выработки навыков работы в команде «на результат».

4. Как следствие культуры проектной работы достигается высокая мотивация студентов, возрастающий интерес к профессии и желание упорно работать для достижения целей учебной программы.

5. Мультидисциплинарность. Все дисциплины, начиная с младших курсов, преподаются в связке друг с другом. Полученные знания немедленно применяются — не обязательно в той же дисциплине. Контроль за эффективной поддержкой мультидисциплинарно-

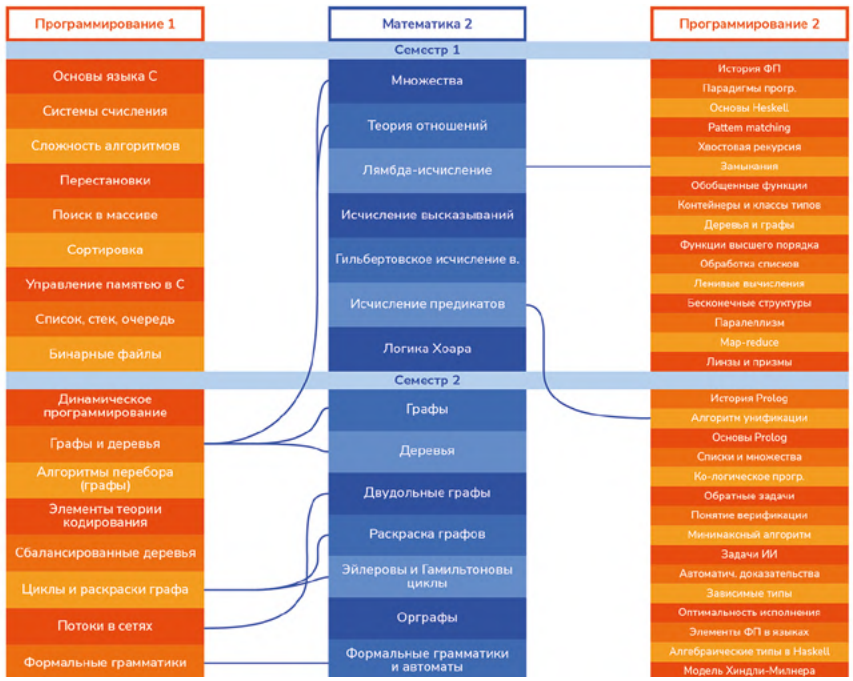


Рис. 1. Схема зависимости разделов профильных дисциплин 1 курса от тем предметной области дискретной математики (курс Математика 2).

сти производится с помощью специальных административных процедур на еженедельной основе.

6. Как следствие (5) — высокая концептуальная слаженность программы, максимизирующая темпы обучения и его результаты.

7. Предмет компьютерных наук — не только традиционное программирование и софтверные технологии. Студенты на данной программе изучают и электронику, в том числе и традиционную аналоговую, и технику измерений и обработки данных. В результате выпускники могут работать в областях, включающих разработку программно-аппаратных комплексов, в задачах интернета вещей, системы управления производством и других. При этом сохраняется полная компетенция в традиционных областях софтверной инженерии.

8. В силу потребности профессии усилена специфическая подготовка по английскому языку, особенно в части составления технической документации и проведения технических обсуждений/презентаций. Часть дисциплин/тематик преподается на английском языке.

Успех данной программы (подготовлено уже два выпуска, и выпускники мгновенно разобраны в престижные места) также основан на применении передовых методов обучения. А именно:

1. В НГУ создан комплекс подготовки студентов по спортивному программированию, содержащий весьма обширную библиотеку упражнений, средства автоматизированной проверки решений, и подготовлены уникальные методики преподавания традиционного программирования на первом курсе. В этой программе они применяются в академических, а не в спортивных целях.

2. В Хартфордшире создан комплекс обучения студентов основам архитектуры ЭВМ, цифровой электроники и низкоуровневому программированию (важнейший компонент в индустриальных приложениях). Создана обширная библиотека упражнений, групповых проектов и средства автоматической проверки решений. Эти материалы переданы НГУ, они поддерживаются и расширяются собственными кадрами.

3. Разработанный в НГУ курс декларативного программирования обладает уникальной педагогической эффективностью. Студенты,

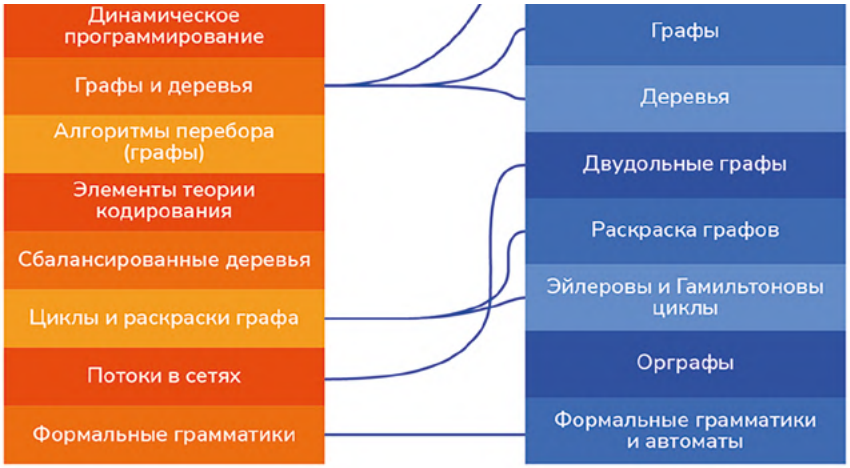


Рис. 2. Зависимость разделов курса Программирование 1 от материала курса Математика 2 во втором семестре.

как правило, мотивированные потребностью практики, изучают этот предмет с удовольствием, хотя он служит в основном для развития кругозора и расширения арсенала технических приемов.

4. Как следствие предыдущего, студенты получают сильный импульс на первом курсе, что помогает им сохранять мотивацию и на втором году обучения, где теоретические дисциплины гораздо сложнее, а проектная и инженерная часть обширнее. Основную мотивирующую роль играет проект второго курса, который курируется специалистами из индустрии, преподающими в НГУ. Таким образом, связь с производством осуществляется с самых младших курсов.

В качестве иллюстрации на рис. 1 представлена зависимость двух разделов профильных дисциплин (Программирование 1 и 2) от тематик, изучаемых в курсе «собирательном» дискретной математики — Математика 2.

На рис. 2 выделены эти зависимости во втором семестре.

Уникальность образовательного продукта, продемонстрированная выше, позволяет говорить о целесообразности его тиражирования и более эффективного использования. Представляется возможным создание на базе НГУ Центра компетенций для обеспечения возможности использования разработанных курсов первой фазы продукта — внедрения концепции в университетах Сибири и Дальнего Востока, а в перспективе и по всей территории РФ, в ЕАЭС и в дружественных странах.

Роль Центра компетенций — заниматься следующими формами деятельности:

1. Курирование использования учебно-методических материалов и автоматизированных систем проверки знаний студентов. Поддержание и расширение курсов первой фазы в части учебных пособий, банков задач, экзаменационных вопросов (тесты для оценки текущей успеваемости), заданий на разработку/проекты с автоматической и полуавтоматической проверкой решений. На этапе организационного становления Центра компетенций на эту работу потребуется выделение средств (со стороны НГУ, профильных министерств, региона или индустриальных партнеров). В дальнейшем Центр компетенций может заниматься экспортом и актуализацией материалов, переводом их на английский язык и др.

2. Подготовка и переподготовка преподавателей. Центр компетенций проводит подготовку преподавателей из системы высшего образования с целью обеспечения их знаниями, умениями и навыками, необходимыми для преподавания разработанных оригинальных учебных дисциплин первой фазы в своих, «домашних», университетах. В течение первых трех лет эта работа оплачивается Центром компетенций из его проектного финансирования, затем на договорной основе.

3. Проведение онлайн-консультаций с преподавателями предлагаемой программы за пределами НГУ. Цель этой работы — дать пользователям уверенность, что если они столкнутся с трудностями в обучении своих студентов, то специалисты НГУ будут на расстоянии звонка от них. Данные консультационные услуги предоставляются из проектного финансирования в первые годы функционирования Центра компетенций, затем на договорной основе.

4. Предоставление удаленного доступа к автоматизированным системам проверки выполнения заданий, включая ведение практических занятий, мониторинг решения задач, получение отчетов об успеваемости и т.п. При этом сами системы автоматической проверки и назначения задач не экспортируются заказчикам, а предоставляются в виде доступа к ним, вначале бесплатного, а после нескольких лет на договорной основе, что послужит закреплению клиентуры за Центром компетенций.

ПРИЛОЖЕНИЕ. УЧЕБНЫЙ ПЛАН ПЕРВЫХ ДВУХ ЛЕТ ОБУЧЕНИЯ.

Название дисциплины	Трудоемкость в зачетных единицах по семестрам			
	1-й год		2-й год	
	1-й семестр	2-й семестр	3-й семестр	4-й семестр
История	3	1		
Иностранный язык	3	3	3	
Основы культуры речи	2			
Введение в алгебру и анализ	5	5		
Цифровые платформы	4	5		
Введение в дискретную математику и математическую логику	5	6		
Императивное программирование	5	6		
Декларативное программирование	3	4		
Операционные системы			4	
Объектно-ориентированное программирование			4	4
Разработка программно-аппаратного комплекса для решения научных и прикладных задач (групповой проект)			3	3
Модели вычислений			3	4
Теория вероятностей и математическая статистика			3	3
Дифференциальные уравнения и теория функций комплексного переменного			4	
Введение в искусственный интеллект			4	
Введение в аналоговую электронику и технику измерений				3
Введение в компьютерные сети				4
Встроенные цифровые системы управления				3
Теория параллелизма				4
Деловой английский язык				2
Философия				2

1

NODE 01
NODE 04
NODE 05
BLOCK 01
NODE 03
BLOCK 01
NODE 01
NODE 02

Актуальные приложения науки

24 Программный комплекс USPARS для решения систем линейных алгебраических уравнений в задачах моделирования

29 Unipro UGENE — свободный продукт для визуализации, анализа и обработки данных в молекулярной биологии и биоинформатике

34 Программный комплекс для разработки интеллектуальных систем поддержки принятия решений в абоформализованных областях

39 Технология прогноза урожайности картофеля с использованием дистанционного зондирования

41 Моделирование распространения эпидемий с учетом экономической и социальной ситуации в регионе

46 Клеточно-автоматная модель динамических процессов в газоугольном растворе

52 Оптические системы панорамной диагностики потоков «ПОЛИС»

56 Оптико-информационная диагностика для энергетики, промышленности и транспорта

ПРОГРАММНЫЙ КОМПЛЕКС USPARS
ДЛЯ РЕШЕНИЯ СИСТЕМ ЛИНЕЙНЫХ АЛГЕБРАИЧЕСКИХ УРАВНЕНИЙ В ЗАДАЧАХ МОДЕЛИРОВАНИЯ

УНИПРО

Общество с ограниченной ответственностью Новосибирский центр информационных технологий «УНИПРО»
630090, г. Новосибирск,
ул. Николаева 11/5, оф. 705

Костин Виктор Иванович
кандидат физико-математических наук, научный руководитель проекта USPARS, ведущий инженер НЦИТ «УНИПРО», старший научный сотрудник ИНГГ СО РАН
vkostin@unipro.ru

Соловьев Сергей Александрович
кандидат физико-математических наук, ведущий инженер НЦИТ «УНИПРО», старший научный сотрудник ИНГГ СО РАН
ssoloviev@unipro.ru

Голосов Иван Сергеевич
заслуженный конструктор РФ, директор НЦИТ «УНИПРО»
тел.: (383) 373-24-63, marketing@unipro.ru

Данилова Юлия Эдуардовна
кандидат физико-математических наук, заместитель директора НЦИТ «УНИПРО»
тел.: (383) 373-24-63, marketing@unipro.ru

Семенов Александр Леонидович
кандидат физико-математических наук, руководитель сектора НЦИТ «УНИПРО»
asemenov@unipro.ru

Быков Дмитрий Олегович
старший инженер НЦИТ «УНИПРО»
dbykov@unipro.ru

Черепанов Михаил Евгеньевич
ведущий инженер НЦИТ «УНИПРО»
mcherepanov@unipro.ru

Новый решатель разреженных СЛАУ с широким функционалом, высокой производительностью и поддержкой разных вычислительных архитектур позволит оптимизировать вычисления в различных научно-технических областях.

Программный комплекс (ПК) USPARS предназначен для решения систем линейных алгебраических уравнения (СЛАУ) с большими разреженными матрицами коэффициентов. Такие СЛАУ возникают при решении научных и прикладных задач численного моделирования физических процессов (теплоперенос, акустические, упругие и электромагнитные волновые процессы), в алгоритмах математического моделирования в различных сферах (экономика, логистика, медицина), при проведении инженерных расчетов в системах CAD/CAE (в задачах стационарной прочности, тепловыделения в многослойных электронных схемах и других).

ПК может использоваться как независимая вычислительная программа или как вычислительное ядро в прикладных ИТ-системах. Для отечественных пользователей USPARS может стать хорошей заменой зарубежных решателей СЛАУ.

Основные характеристики

Для численного решения больших СЛАУ с разреженными матрицами коэффициентов, у которых число строк превышает десятки тысяч,

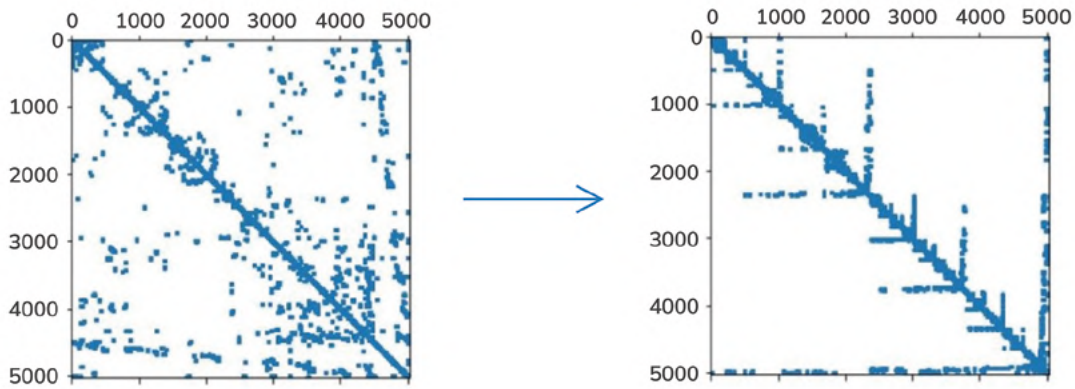


Рис. 1. Визуальный портрет матрицы с размерностью 5000 (25 млн коэффициентов) до и после этапа переупорядочивания элементов.

современная вычислительная линейная алгебра предлагает так называемые прямые и итерационные методы. У каждого из двух подходов имеются свои достоинства и недостатки.

В USPARS используется прямой метод Гаусса решения разреженных СЛАУ, в котором матрица коэффициентов раскладывается в произведение треугольных факторов (так называемая факторизация), после чего задача оказывается сведенной к решению СЛАУ с треугольными матрицами коэффициентов. Основным достоинством метода Гаусса является меньшая, по сравнению с итерационными методами, чувствительность вычислений к величине числа обусловленности матрицы коэффициентов. Это свойство формально расширяет круг решаемых СЛАУ. Кроме того, имеется возможность переиспользования результатов факторизации для решения СЛАУ с той же самой матрицей коэффициентов, но другими векторами правой части.

Для хранения получаемых в результате факторизации треугольных факторов часто требуется значительно больше памяти, чем для исходной матрицы, и это является существенным осложнением при использовании метода Гаусса. Объем требуемой памяти зависит от структуры ненулевых элементов матрицы коэффициентов. Однако эффективные методы переупорядочивания матрицы коэффициентов, примененные перед факторизацией, позволяют существенно снизить требования к объему доступной оперативной памяти. Для переупорядочивания матрицы с целью сокращения объема требуемой памяти в USPARS применяется собственная оригинальная разработка.

ПК USPARS предоставляет возможность оперировать с данными вещественного и комплексного типов одинарной и двойной точности. Опционально для улучшения точности решения можно воспользоваться итерационным уточнением.

Ряд специальных мер предпринят для различного рода оптимизаций:

- компонента переупорядочивания матрицы служит для оптимизации расходов оперативной памяти и числа арифметических операций, требующихся для получения решения;
- применены оптимизированные алгоритмы факторизации больших разреженных матриц;
- возможно использование на многопроцессорных вычислительных системах с общей памятью; распараллеливание кодов выполнено с использованием открытого стандарта OpenMP.

Для удобства пользователя в пакете предоставляется Scipy-совместимый python-интерфейс с простым синтаксисом запросов, автоматическим контролем памяти, поддержкой стандартных Scipy-визуализаций (см. рис. 1) и Scipy CSR формата матриц.

Функциональность пакета независима от стороннего лицензионного программного обеспечения. Также обеспечена поддержка трех процессорных архитектур — x86, ARM и Эльбрус.

Наличие опытных образцов, результатов их испытаний

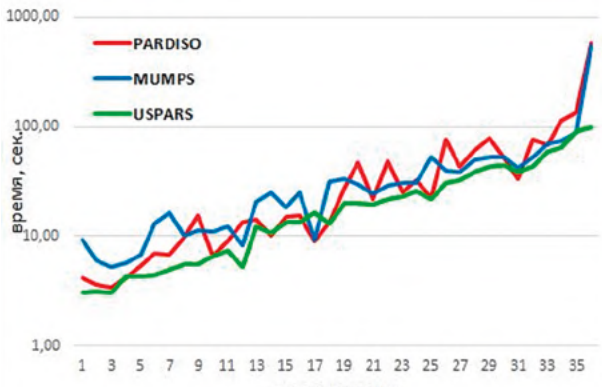
Пакет USPARS представляет собой готовый программный продукт с подробной пользовательской документацией. Он непрерывно развивается, появляются новые функции и улучшаются уже имеющиеся. Обновленные версии пакета выпускаются несколько раз в год.

Тщательно проработанный API обеспечивает удобство как непосредственного использования функций пакета, так и включения его в качестве вычислительного ядра в инженерные системы.

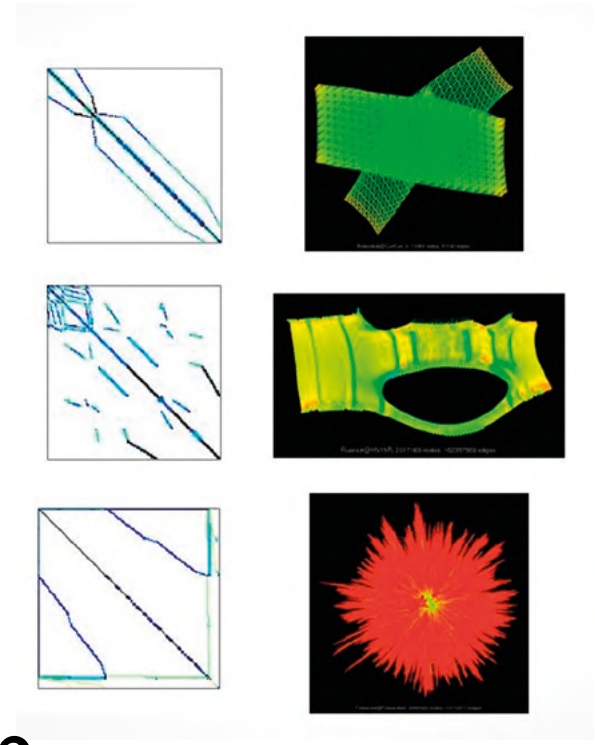
В конце 2022 года пакет успешно введен в эксплуатацию в одном из ведущих российских научно-производственных центров ГК «Росатом».

Общее время, требующееся на решение той или иной СЛАУ, является важной характеристикой потребительских свойств пакета. Это время варьируется в зависимости от характеристики используемого «железа» (версия процессора, его тактовая частота, число ядер, объем оперативной памяти и т.д.), порядка матрицы, структуры ее ненулевых элементов, эффективности распараллеливания кода и множества других факторов.

На рис. 2 представлены результаты сравнения последней версии USPARS и свежих версий двух популярных зарубежных пакетов на сравнительно небольшой, но представительной выборке из 36 матриц, взятых с ресурса SuiteSparse Matrix Collection. Порядки матриц в этой выборке варьируются от 14 тыс. до 5.5 млн, а число ненулевых элементов — от 800 тыс. до 110 млн. Матрицы сравнительно неплохо обусловлены — их числа обусловленности лежат в интервале от 10^2 до 10^{14}. Структуры матриц существенно различаются, что показано на примере трех из них в двух визуальных представлениях на рис. 3.



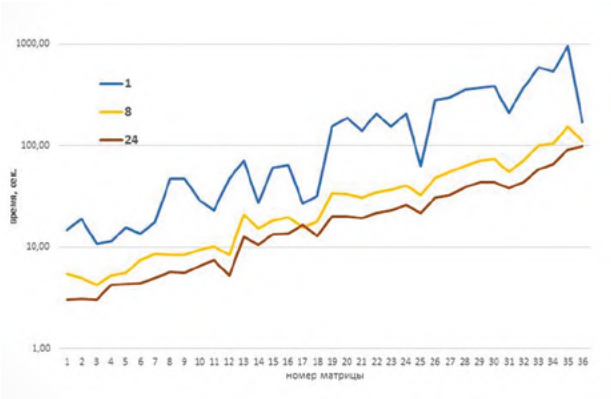
2



3

Визуально USPARS существенно опережает конкурентов (на этой выборке матриц) — зеленая линия, представляющая вычислительные времена USPARS, всюду за редкими исключениями, лежит ниже двух других. Отношение вычислительных времен, потраченных двумя различными пакетами для решения одной и той же СЛАУ (конкретной матрицы), характеризует превосходство второго над первым. Среднее геометрическое ω этих отношений по всей выборке матриц дает интегральную характеристику сравнения двух пакетов. Для результатов, представленных на рис. 2, получено ω (PARDISO, USPARS) =1.4, ω (MUMPS, USPARS)=1.6, что подтверждает сформулированный выше вывод о превосходстве USPARS на данной выборке матриц.

Программный код пакета USPARS распараллелен с использованием OpenMP. Это означает, что при работе на вычислительной системе с общей памятью, построенной с использованием многоядерных процессоров, имеется возможность задействовать несколько потоков и при увеличении числа потоков ожидать уменьшения времени решения СЛАУ. Отношение t_1 (A)/ t_p (A) времени решения t_1 СЛАУ с матрицей коэффициентов A



4

на вычислительной системе с задействованным одним потоком ко времени решения t_p (A) той же СЛАУ на той же системе с задействованными p потоками носит название масштабируемости программы. В том случае, если отношение t_1 (A)/ t_p (A) равно p, масштабируемость называют идеальной, но достичь ее практически невозможно.

Для упомянутой выше выборки матриц на рис. 4 представлены три графика вычислительного времени — для 1, 8 и 24 потоков, соответствующих 1, 8 и 24 ядрам сервера упомянутого на рис.2. При этом, как и раньше, по горизонтальной оси стоят номера матриц из выборки. При p=8 минимальное значение масштабируемости равно 1.55, максимальное — 6.29. Соответственно, при p=24 минимальное значение масштабируемости равно 1.65, максимальное — 10.51. Если вычислить среднее геометрическое отношений t_p (A)/ t_1(A) по заданной выборке, получим 3.62 для p=8 и 5.48 для p=24. Видно, что масштабируемость USPARS на матрицах выборки не идеальна, и есть потенциал для ее дальнейшего улучшения. Однако основным критерием все же является производительность пакета, которая, как показано выше, уже находится на уровне конкурентов.

Уровень готовности технологии (TRL) и дорожная карта

USPARS имеет уровень готовности технологии УГТ 8 согласно ГОСТ Р 58048–2017. Это означает, что все функциональные возможности пакета успешно продемонстрированы в смоделированной эксплуатационной среде как у нас, так и у заказчика; проведен полный цикл тестирования и отладки; подготовлен полный пакет технической и пользовательской документации.

Команда имеет план развития продукта на несколько лет вперед. Дорожная карта определяется как собственным видением разработчиков, так и пожеланиями со стороны заказчиков. Одна из ближайших задач — разработка кластерной MPI версии пакета USPARS для работы на вычислительных системах с распределенной памятью. Также в планах команды расширение USPARS на другие процессорные архитектуры.

Преимущества и уникальность

USPARS решает задачу импортозамещения зарубежных решателей СЛАУ, находясь уже на одном уровне с ними по параметрам решаемых задач и качеству результатов. В течение ближайших 2–3 лет пакет имеет хорошие шансы перегнать некоторых конкурентов за счет развития нового функционала и увеличения производительности вычислений. Для активно развивающегося отечественного инженерного ПО это важный и необходимый компонент. На зарубежном рынке пакет также может привлечь интерес и получить заказчиков.

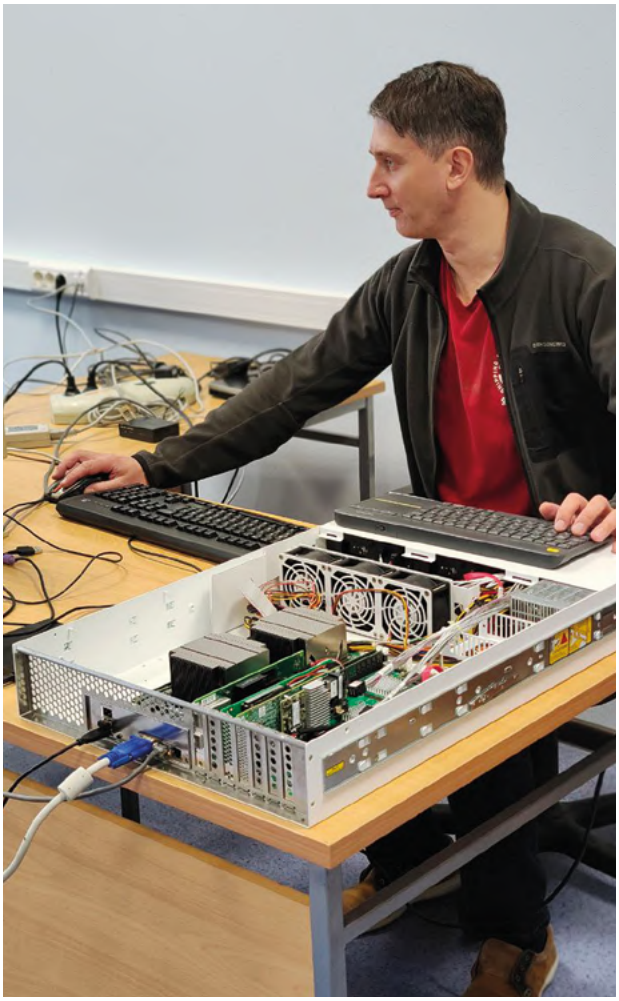
Продукт сочетает в себе классический функционал прямых решателей и ряд авторских алгоритмических оптимизаций, по сути ноу-хау, обеспечивающее высокую производительность вычислений.

В начале 2022 года на USPARS получено Свидетельство о государственной регистрации программы для ЭВМ N2022614559 от 5 марта 2022 г.

2. Время расчета тремя решателями — USPARS, PARDISO, MUMPS — выборки матриц из тестовой сюиты SuiteSparse Matrix Collection на сервере 2CPU Intel(R) Xeon(R) CPU E5–2680 v3 @ 2.50GHz (12 core, HT disabled), 256 GB RAM, OS Ubuntu 20.04.3 LTS

3. Портреты (слева) и графы (справа) трех матриц из коллекции SuiteSparse Matrix Collection http://sparse.tamu.edu/: моделирование полупроводников (Bodendiek/CurlCurl_0), задачи гидродинамики (Fluorem/HV15R), схемотехническое моделирование (Freescale/Freescale2).

4. Зависимость времени решения от количества (1, 8 или 24) используемых потоков на сервере Intel.



Ведущий разработчик Соловьев Сергей Александрович за очередным запуском пакета USPARS

Недавно продукт зарегистрирован в Реестре российского ПО, номер реестровой записи 18945 от 5.09.2023.

История и аналоги

Отрасль математических решателей включает в себя различные направления. Одно из них — программные пакеты для решения разреженных СЛАУ, существует уже более шестидесяти лет. Сейчас в мире известно порядка десятка широко используемых решателей разреженных СЛАУ, коммерческих и open-source. Они стремительно эволюционируют вследствие растущих потребностей практических задач, развития компьютерной техники и сложных прикладных программных пакетов.

В 2020 году НЦИТ «УНИПРО» собрал в новый проект USPARS коллектив с серьезным научно-практическим заделом в разработке инженерно-математических программных пакетов, высокопроизводительных эффективных алгоритмов и разработке различных алгоритмов и подходов для решения СЛАУ. И многолетний опыт самой ИТ-компании в разработке наукоемкого и коммерческого ПО тоже немаловажен. Получилось удачное сочетание фундаментальных научных знаний и опыта прикладных программных разработок.

Широко известными аналогами USPARS являются зарубежные пакеты MKL PARDISO, MUMPS, PARDISO, SuperLU, TAUCS, UMFPACK.

Российские решатели есть, но они малоизвестны. Это объясняется, вероятно, использованием в узком «коридоре», от конкретного разработчика конкретному заказчику, или тем, что они «зашиты» внутри готовых ИТ-систем.

Предполагаемый интерес для внедрения

В РФ есть несколько сильных команд, разрабатывающих инженерное ПО. С одной из них мы сотрудничаем уже второй год.

Для объединения и корректировки усилий летом 2021 года был создан отечественный консорциум разработчиков инженерных систем классов CAD/CAE. В консорциуме ведется работа по интеграции ПО классов CAD и CAE на вычислительную платформу Росатома «Логос Платформа». Надеемся, что наш универсальный решатель станет составной частью этой платформы и будет востребован в различных прикладных научно-технических и индустриальных проектах.

Команда готова развивать продукт в интересах разных индустриальных и научно-технических партнеров. Мы можем поставлять USPARS как самостоятельный лицензионный продукт и готовы развивать его под заказ, удовлетворяя нужды конкретного заказчика. Оба пути для нас привычны. Техническую поддержку мы обеспечиваем в любом варианте.

UNIPRO UGENE — СВОБОДНЫЙ ПРОДУКТ

ДЛЯ ВИЗУАЛИЗАЦИИ, АНАЛИЗА И ОБРАБОТКИ ДАННЫХ В МОЛЕКУЛЯРНОЙ БИОЛОГИИ И БИОИНФОРМАТИКЕ

УНИПРО

Общество с ограниченной ответственностью Новосибирский центр информационных технологий «УНИПРО»
630090, г. Новосибирск,
ул. Николаева 11/5, оф. 705

Сухомлинов Дмитрий Игоревич
ведущий инженер
тел.: (383) 373-24-63, ugene@unipro.ru

Данилова Юлия Эдуардовна
кандидат физико-математических наук, заместитель директора по маркетингу
тел.: (383) 373-24-63, marketing@unipro.ru

Сухомлинов Дмитрий Игоревич, разработчик продукта

Данилова Юлия Эдуардовна, маркетолог продукта

В течение последних десятилетий цифровизация биологических исследований превратилась из узконаправленной области в глобальную — всё больше и больше экспериментов перемещается из лабораторий в компьютеры. При этом количество инструментов, предназначенных для того или иного вида анализа, растет с такой скоростью, что, даже будучи специалистом, трудно отследить их появление. С другой стороны, цифровые биологические данные должны иметь визуально привлекательный и удобный внешний вид, позволяющий с комфортом осуществлять работу, не отвлекаясь на их сложную внутреннюю структуру.

Unipro UGENE является инструментом для повседневных молекулярно-биологических процедур, объединяющий в себе все указанные требования. UGENE предоставляет удобный графический интерфейс для работы со всеми возможными типами биоинформатических данных — генетическими последовательностями, множественными выравниваниями, прочтениями высокопроизводительного секвенирования и их сборками, прочтениями по Сэнгеру, филогенетическими деревьями и 3D-моделями белков. Каждый из представленных видов данных может быть проанализирован и обработан десятками различных способов в зависимости от поставленных задач — аннотирование, поиск функциональных элементов, подбор праймеров, клонирование, выравнивание набора последовательностей друг на друга или на другую последовательность, построение филогенетических деревьев и т.д. (рис. 1)

Помимо собственных алгоритмов и визуальных решений UGENE решает другую важную задачу — он, как платформа, включает в себя внешние биоинформатические инструменты разных направлений, каковых в наши дни существует великое множество. Для их запуска в нем создаются удобные графические интерфейсы с подробными

инструкциями по всем шагам работы. В данный момент в состав UGENE входит более двадцати внешних инструментов, например BLAST, ClustalW, Primer3, FastQC и многие другие. Более того, UGENE позволяет интегрировать любую функциональную программу — будь то полноценный инструмент или небольшой скрипт — в качестве элемента вычислительного конвейера.

Такие вычислительные конвейеры представляют собой схемы, собранные из отдельных элементов, каждый из которых являет собой определенную вычислительную задачу. Пример показан на рис. 2. В UGENE подобные схемы создаются, отлаживаются и выполняются с помощью «Дизайнера вычислительных схем». Данная технология позволяет совершать наборы однотипных расчетов на большом количестве данных, а также организовывать подобные расчеты в виде серверных решений.

Кроме непосредственно разработки продукта наша команда участвует в различных научных мероприятиях. Например, на конференциях Bioinformatics of Genome Regulation and Structure/Systems Biology и OpenBio мы много лет представляем доклады по новым биоинформационным разработкам. На школах молодых ученых в качестве учителей делимся знаниями о том, как по-разному можно анализировать данные с помощью UGENE. Хороший пример — это ежегодная международная школа- конференция молодых вирусологов The Virus Evolution and Molecular Epidemiology (VEME). Несколько лет назад мы разработали курс «Практическая биоинформатика», материалы которого используются в преподавании для магистрантов НГУ. И гордимся тем, что студенческий трек профиля «Геномное редактирование» олимпиады НТИ с момента своего старта в 2019 году использует UGENE как основной инструмент для решения компьютерных задач с генетическими последовательностями — уже несколько тысяч студентов и старшеклассников получили практические навыки работы в нем.

Команда постоянно взаимодействует с пользователями со всего мира (рис. 3), помогая решать технические и концептуальные вопросы: по электронной почте, на форуме, в соцсетях.

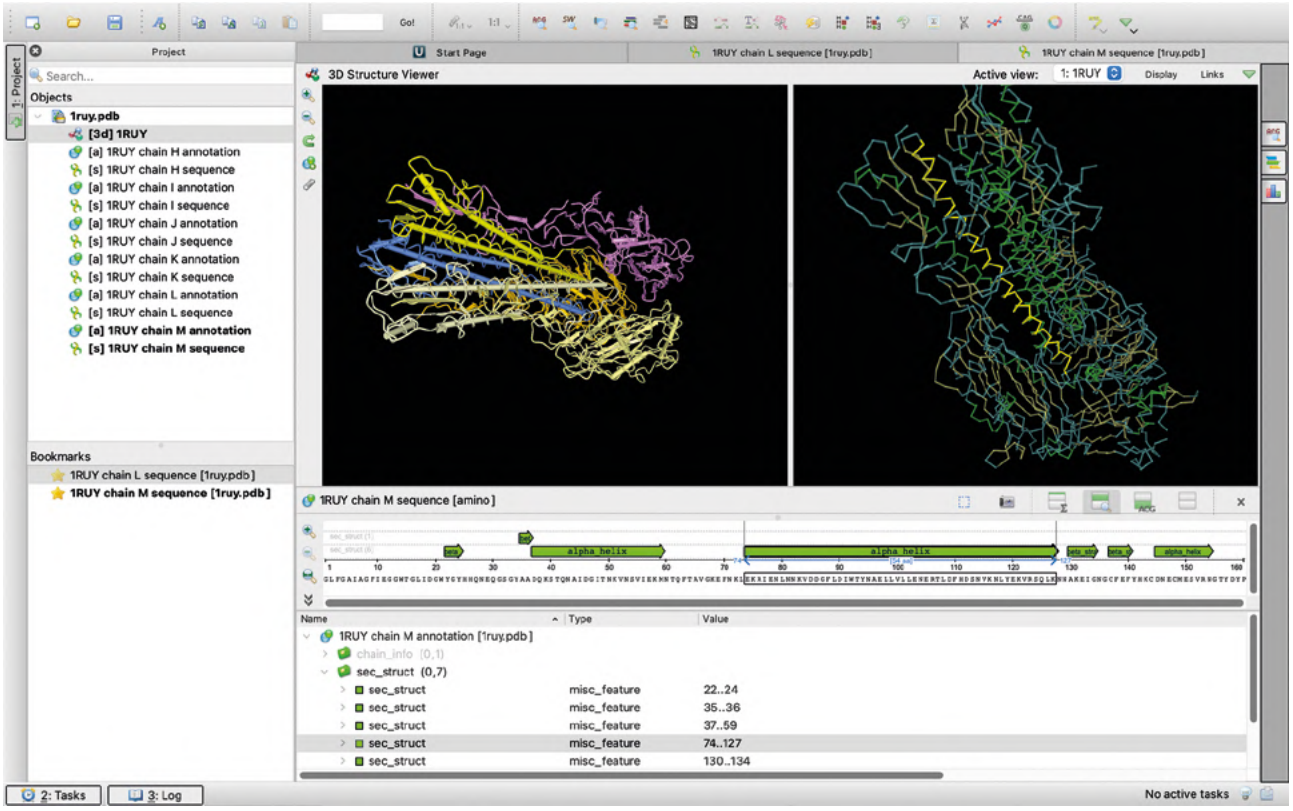


Рис. 1. Пример интерфейса UGENE. Набор белковых последовательностей и визуализация их 3D-структур.

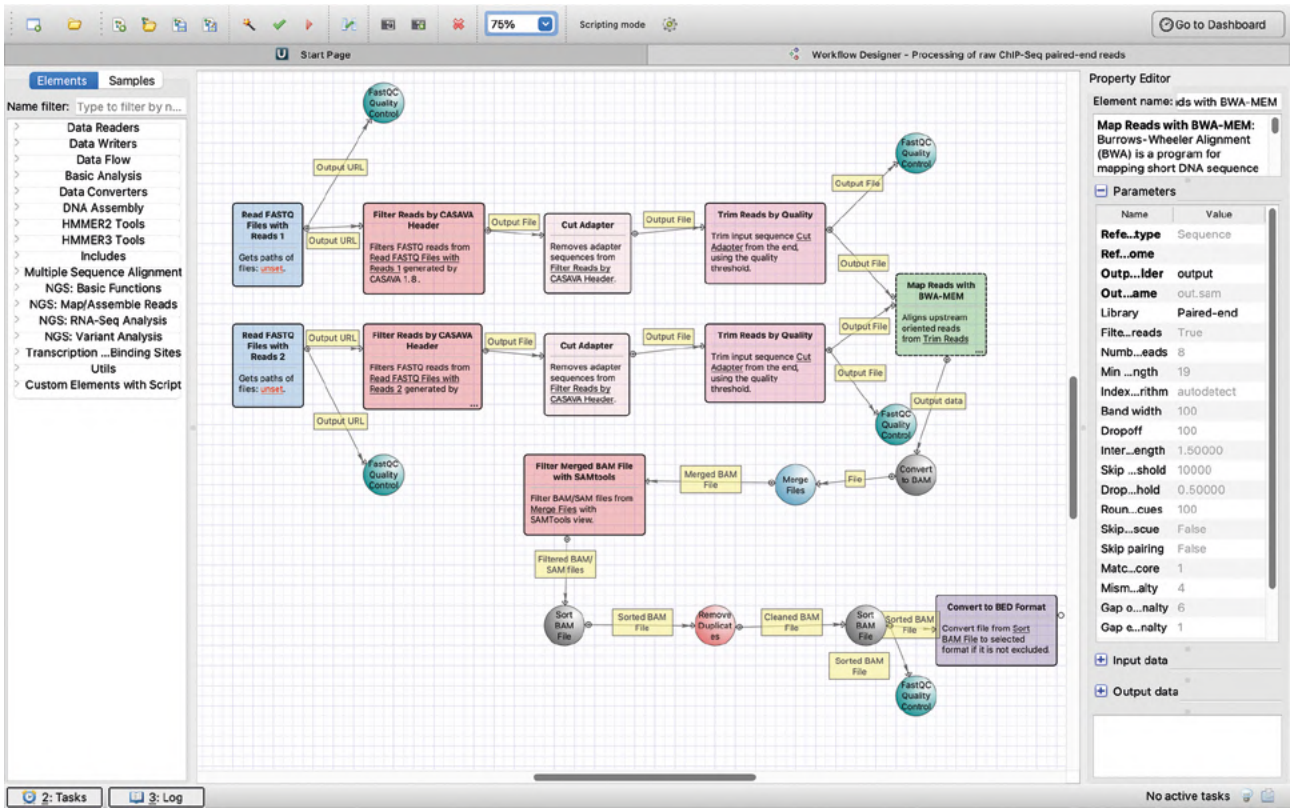


Рис. 2. Дизайнер вычислительных схем в UGENE с готовым вычислительным конвейером для анализа и обработки «сырых» данных секвенирования.

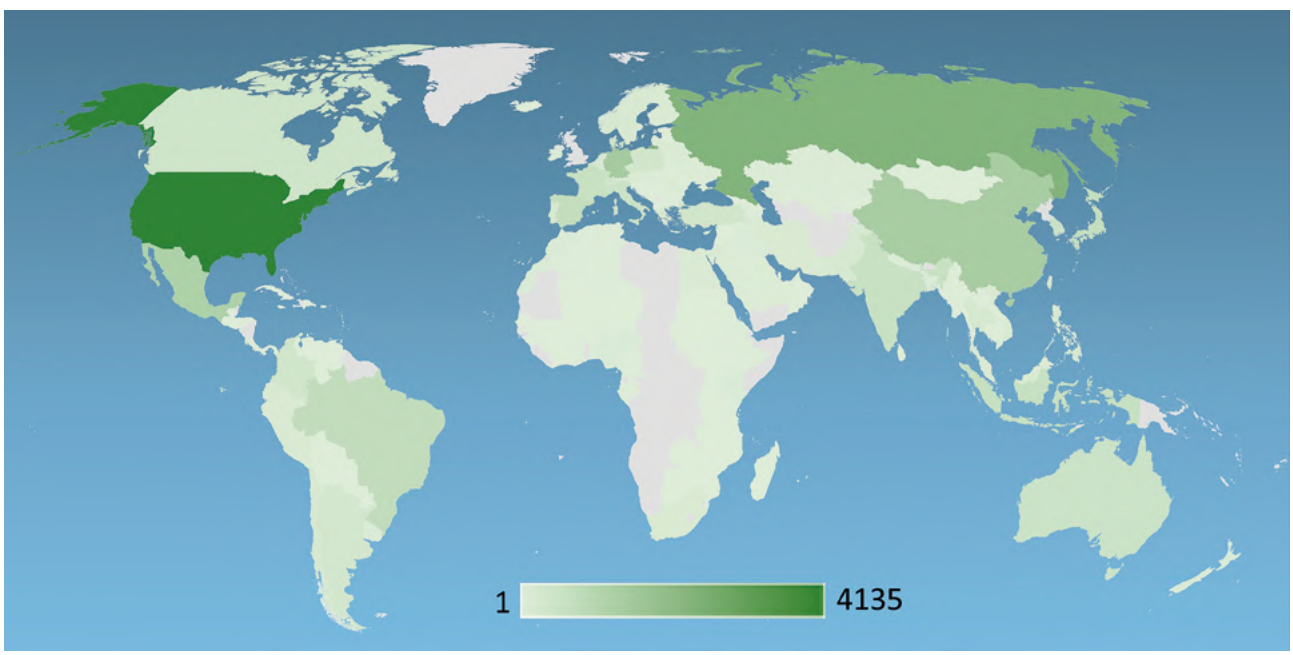


Рис. 3. Карта, показывающая распределение пользователей UGENE по странам (статистика за 2022 год). Как видно, первое место по этому показателю занимает США, второе — Россия. Третье место с относительно небольшим отрывом друг от друга делят Китай, Германия и Мексика.

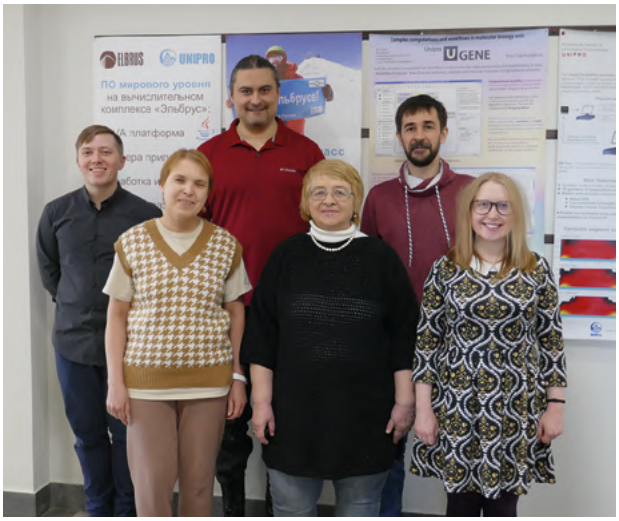


Рис. 4. Команда разработки и продвижения UGENE. Слева направо: Сухомлинов Дмитрий, Данилова Юлия, Распутин Кирилл, Самойленко Светлана, Варламов Алексей, Бисерова Эвелина.

Рис. 5. Кандров Денис, один из ветеранов-разработчиков UGENE.

Много лет разработчики ведут обучающий подкаст на Youtube, постоянно обновляют документацию продукта. По запросам организаций проводится онлайн- и офлайн-обучение для их сотрудников.

Продукт распространяется бесплатно по лицензии GPL, что делает его максимально доступным для всех пользователей. Нужно просто зайти на сайт ugene.net и скачать нужную вам версию.

Уровень готовности технологии

Unipro UGENE существует уже более 15 лет — первая версия увидела свет в 2008 году. За это время было выпущено 48 релизов, каждый из которых постепенно развивал и дополнял программу; написано семь статей, процитированных, в общей сумме порядка 2700 раз. Также на базе UGENE разработан ряд продуктов, решающих частные задачи в сфере биоинформатики. Только за 2022 год приложением воспользовалось более 35 тысяч пользователей, запустив Unipro UGENE порядка 370 тысяч раз. С учетом вышесказанного можно утверждать, что Technology Readiness Level (TRL) продукта равен 9.

UGENE имеет Свидетельство о государственной регистрации программы для ЭВМ N 2016615455 от 25 мая 2016 года и состоит в Реестре российского ПО, номер реестровой записи 2883 от 14.03.2017.

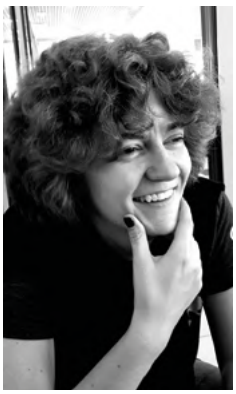
Преимущества и уникальность

Unipro UGENE является аналогом швейцарского ножа в биоинформатике, объединяя удобный и привлекательный графический интерфейс с возможностью работы с любыми видами данных, а также множество алгоритмов и методов, позволяющих обрабатывать, анализировать и улучшать эти данные. Это выгодно отличает UGENE от многих узкоспециализированных биоинформационных инструментов, призванных решать только частные задачи.

Несмотря на наличие ряда зарубежных аналогов, UGENE является единственной отечественной разработкой с подобным функционалом, что гарантирует наличие конкурентоспособного продукта в данной сфере вне зависимости от налагаемых ограничений на иностранные программные продукты. А политика свободного распространения UGENE делает его доступным для всех.

История

Работы в области биоинформатики начались в НЦИТ «УНИПРО» в 2003 году с небольшого продукта для сравнения двух генетических последовательностей — DotPlot. Позднее он вошел в состав первого большого продукта, получившего имя Genome Browser, который предназначался для работы с отдельными геномами.



Альгаер Юлия,
менеджер проекта
UGENE

Личный комментарий от руководителя проекта:

«Проработав в проекте 12 лет, можно рассказать много разных историй, как успешных, так и тех, что принесли ценный опыт. Наша модель работы является уникальной на рынке и позволяет минимизировать риски для клиента: код проекта открыт, не зависит от политических изменений, UGENE не требует лицензионных отчислений и может использоваться как в небольших, так и в крупных огрганизациях. Главным активом UGENE является команда. За 15 лет существования через проект прошли десятки талантливых разработчиков, и мы гордимся тем, что смогли сохранить ядро команды, а значит и опыт. И этот опыт сегодня доступен всем, кто работает с нами. Присоединяйтесь!».

В 2008 году этот проект трансформировался в основной биоинформационный проект компании — Unipro UGENE, который унаследовал возможности предшественника и существенно расширил их.

Создавая UGENE, команда УНИПРО ориентировалась в первую очередь на конкурентоспособность в сравнении с зарубежными аналогами. И это удалось — UGENE пережил несколько иностранных продуктов, на которые равнялся еще на этапе своего создания (из наиболее крупных примеров — Vector NTI, чья разработка была прекращена в 2019 году). Поддерживать высокую востребованность в мировом биологическом сообществе помогает сотрудничество с экспертами международного уровня.

Сотрудничество

Unipro UGENE представляет интерес для исследователей и организаций, работающих в молекулярной биологии и биоинформатике, например, в таких направлениях, как секвенирование, геномное редактирование, метагеномика и т.д. Он востребован как в академической среде и образовательной сфере, так и в прикладной биотехнологии и биомедицине.

За многие годы налажено большое количество контактов с научными институтами, университетами и частными исследовательскими компаниями. Одними из ближайших наших соседей и коллег были и остаются ИЦиГ СО РАН и ИХБФМ СО РАН. Ряд функций в составе UGENE появились благодаря сотрудничеству с этими институтами. В числе заказчиков значатся как российские организаци, например новосибирский «Вектор-Бест», Санкт- Петербургская «Секвойя Дженетикс», так и зарубежные, например европейский консорциум VIROGENESIS.Биоинформационная команда УНИПРО готова и дальше сотрудничать с различными партнерами как в РФ, так и за рубежом для создания нового и улучшения текущего функционала Unipro UGENE.

Литература

1. Okonechnikov K, Golosova O, Fursov M, the UGENE team. Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics 2012 28: 1166–1167. doi:10.1093/bioinformatics/bts091

2. Golosova O, Henderson R, Vaskin Y, Gabrielian A, Grekhov G, Nagarajan V, Oler AJ, Quiñones M, Hurt D, Fursov M, Huyen Y. Unipro UGENE NGS pipelines and components for variant calling, RNA-seq and ChIP-seq data analyses. PeerJ 2014 2: e644. doi:10.7717/peerj.644

3. Rose R, Golosova O, Sukhomlinov D, Tiunov A, Prosperi M. Flexible design of multiple metagenomics classification pipelines with UGENE. Bioinformatics, bty901, 2018/10/25. doi:10.1093/bioinformatics/bty901

ПРОГРАММНЫЙ КОМПЛЕКС ДЛЯ РАЗРАБОТКИ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ ПОДДЕРЖКИ

ПРИНЯТИЯ РЕШЕНИЙ В СЛАБОФОРМАЛИЗОВАННЫХ ОБЛАСТЯХ

Институт систем информатики им. А.П. Ершова СО РАН

Федеральное государственное бюджетное учреждение науки Институт систем информатики им. А.П. Ершова Сибирского отделения Российской академии наук
630090, г. Новосибирск,
пр. Академика Лаврентьева, 6

Загорулько Юрий Алексеевич
кандидат технических наук, заведующий лабораторией искусственного интеллекта
тел.: (383) 332-83-59, zagor@iis.nsk.su

Загорулько Галина Борисовна
кандидат технических наук, старший научный сотрудник
тел.: (383) 332-83-59; gal@iis.nsk.su

Серый Алексей Сергеевич
младший научный сотрудник
тел.: (383) 330-79-91, Alexey.Seryj@iis.nsk.su,

Шестаков Владимир Константинович
младший научный сотрудник
тел.: (383) 330-79-91, umanist@gmail.com

Разработчики СПОРА+: Загорулько Галина Борисовна, Шестаков Владимир Константинович.

Сотрудники лаборатории искусственного интеллекта, принимавшие участие в разработке СПОРА+ (слева направо): Шестаков Владимир Константинович, Серый Алексей Сергеевич, Загорулько Юрий Алексеевич, Ахмадеева Ирина Равильевна

Интеллектуальные системы поддержки принятия решений (ИСППР) — это востребованный класс программных систем, которые создаются в разных предметных областях для решения слабоструктурированных проблем, характери-

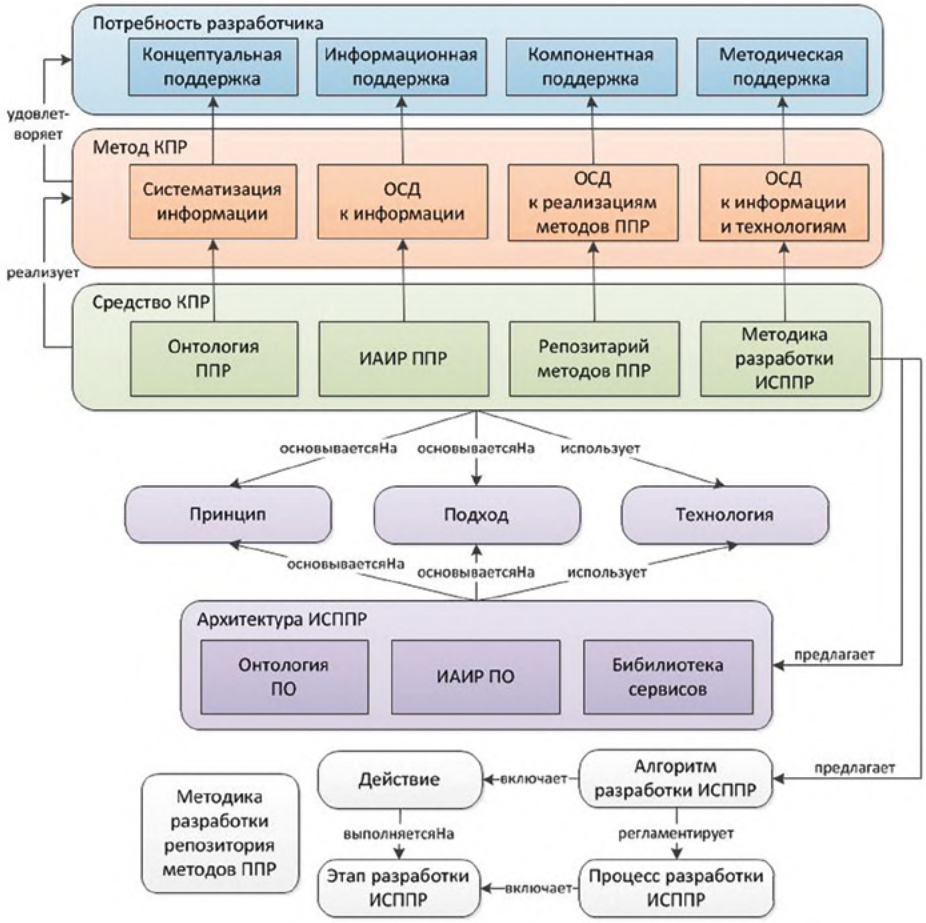


Рис. 1. Модель комплексной поддержки разработчиков ИСППР. Здесь КПР — комплексная поддержка разработчиков, ОСД — организация содержательного доступа, ПО — предметная область.

зующихся отсутствием строгих математических моделей и использованием преимущественно качественной информации в условиях НЕ-факторов (неполноты, неточности, неопределенности и прочее). Современная прикладная ИСППР, как правило, состоит из нескольких взаимосвязанных подсистем: информационной, моделирующей, аналитической, оценивающей и интеллектуальной (экспертной). Эти подсистемы решают разные задачи и используют для этого разные методы. Разработка таких систем — важная задача, решение которой осложняется отсутствием универсальных доступных инструментариев.

В ИСИ СО РАН предложена модель поддержки разработчиков ИСППР (рис. 1) и создан программный комплекс СПОРА+, предназначенный для разработки ИСППР. СПОРА+ учитывает потребности разработчиков и предлагает средства удовлетворения этих потребностей на всех этапах разработки ИСППР, оказывая им концептуальную, информационную, компонентную и методическую поддержку.

В разработке ИСППР участвуют специалисты разного профиля: эксперты той предметной области, для которой создается ИСППР; инженеры знаний, являющиеся специалистами в области представления знаний и поддержки принятия решений; программисты. Все они говорят на своем профессиональном языке, используют свою терминологию. Для их успешной совместной работы, в особенности на начальных этапах разработки, необходима единая система понятий, общий концептуальный базис. СПОРА+ предоставляет в качестве средства концептуальной поддержки разработчиков ИСППР онтологию области знаний «Поддержка принятия решений» (ППР).

Помимо общих представлений об области знаний разработчикам также необходима инфор-

мация о конкретных методах ППР, о классах задач, решаемых этими методами, о возможностях и ограничениях каждого из них. Кроме этого разработчикам должна быть доступна информация об основных этапах принятия решений и используемых на каждом из них методах, а также об инструментальных средствах, реализующих эти методы. В качестве средства информационной поддержки в комплекс СПОРА+ включен информационно-аналитический интернет-ресурс (ИАИР) по ППР, разработанный на основе онтологии ППР. На рис. 2 показана страница этого ресурса с описанием метода недоопределенных вычислений, который является эффективным методом принятия решений в слабоформализованных областях [1].

На этапе реализации ИСППР важную роль играет компонентная поддержка разработчиков.Действительно, возможность выбрать и опробовать готовые программные компоненты, реализующие необходимые методы ППР или организующие интерфейс с пользователями, может существенно облегчить и ускорить процесс создания ИСППР. Средством компонентной поддержки в СПОРА+ является репозиторий методов ППР.

Еще одной важной потребностью разработчиков на этапе реализации является методическая поддержка, включающая описание подходов и принципов разработки ИСППР, средств разработки и методику их применения.

Рис. 2. Информационно-аналитический интернет-ресурс «Поддержка принятия решений».

Все средства, предоставляемые СПОРА+, основаны на хорошо зарекомендовавших себя на практике принципах разработки ИСППР: максимальное использование готовых решений, масштабируемость, доступность, открытость, простота использования, независимость от предметной области, информативность. Соблюдение этих принципов обеспечивается использованием подходов быстрого прототипирования и гибкой разработки, а также онтологического, сервис-ориентированного и каркасного подходов [2].

При реализации системы СПОРА+ были использованы сторонние и авторские технологии и программные разработки: технология Semantic Web [3]; технология разработки интеллектуальных научных интернет-ресурсов (ИНИР) [4], ранее созданная в нашем институте, и предоставляющая оболочку ИНИР и ее базовое наполнение в качестве каркаса будущей ИСППР; сервисы, реализующие ряд интеллектуальных методов ППР. Средства, предлагаемые СПОРА+, заложенные в нее принципы и используемые подходы в совокупности обеспечивают широкий набор функциональных возможностей, позволяющих быстро строить и легко использовать на практике ИСППР в различных предметных областях.

Области возможного использования

Программный комплекс «СПОРА+» может быть использован как:

- инструментарий разработки ИСППР в различных слабоформализованных предметных областях — медицине, юриспруденции, образовании, техническом или медицинском мониторинге и диагностике, промышленности, политике, социальной сфере;
- интеллектуальный справочник по области знаний «Поддержка принятия решений»;
- учебный полигон («песочница») для ознакомления с методами поддержки принятия решений и их апробации.

Степень готовности разработки к практическому применению

Программный комплекс СПОРА+ был успешно использован для разработки ИСППР в технических областях и медицинской диагностике. На рис. 3 показан пользовательский интерфейс ИСППР «Угрозы энергетической безопасности» [5], роль которого играет информационно-аналитический интернет-ресурс. На представленной странице ресурса показано описание практической задачи — угроза аварии у производителя энергоресурса. Пользователь может посмотреть текстовое описание этой задачи, ее формальное описание в виде недоопределенной вычислительной модели, которую можно передать соответствующему сервису, и получить результаты. Ссылка на этот сервис также представлена на данной странице.

Входящая в состав СПОРА+ оболочка информационной подсистемы ИСППР была использована для создания целого ряда информационно-аналитических интернет-ресурсов.

Использование программного комплекса СПОРА+ позволит упростить и ускорить разработку ИСППР в слабоформализованных предметных областях.

Контекст

Универсальных средств разработки ИСППР в слабоформализованных предметных областях (СФПО) в свободном доступе нет. Имеющиеся программные средства либо недоступны из-за высокой стоимости (как, например, система G2), либо не предназначены для СФПО (базы данных и электронные таблицы со встроенными средствами OLAP), либо ориентированы на определенные методы принятия решений (системы Colibri, Clips, Netica). Система СПОРА+ основана на интеграции авторских и имеющихся в свободном доступе средств разработки ИСППР и методов поддержки принятия решений. Она предоставляет информацию об этих средствах и реализациях методов, обеспечивает к ним доступ, а также предоставляет «каркас» будущей ИСППР и методику ее разработки. Таким образом, СПОРА+ агрегирует достоинства доступных программных разработок как собственных, так и сторонних авторов, обеспечивая уникальный синергетический эффект от их совместного использования, позволяющий быстро строить ИСППР в различных СФПО.

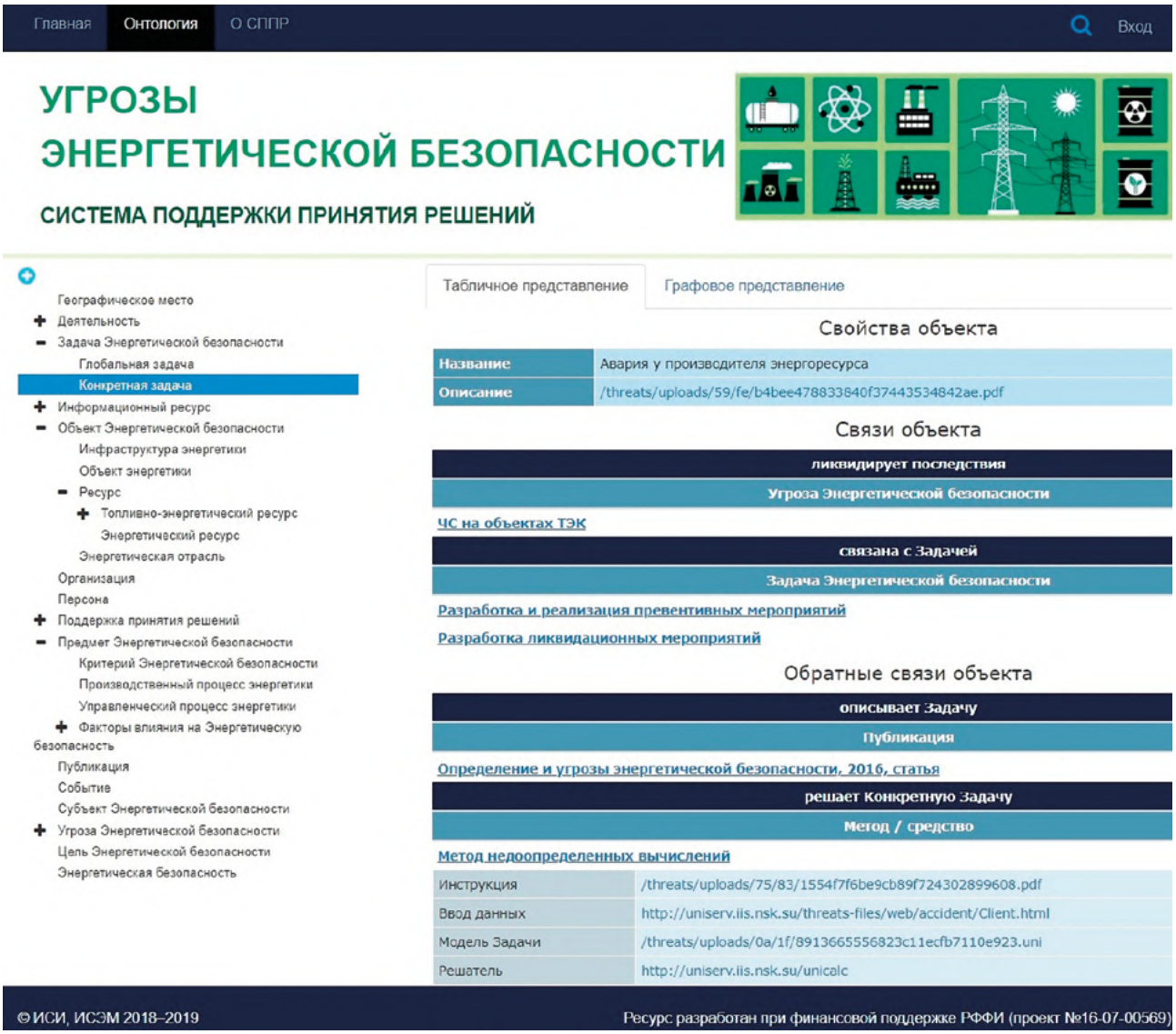


Рис. 3. ИСППР «Угрозы энергетической безопасности».

Сведения о патентоспособности и патентной защите разработки

Программный комплекс СПОРА+ зарегистрирован как программа для ЭВМ в Федеральной службе по интеллектуальной собственности (Роспатент): Свидетельство № 2019619698 от 23 июля 2019 г.

Литература:

1. Загорулько Г. Б., Сидоров В. А. Метод недоопределенных вычислений как средство поддержки принятия решений в слабоформализованных предметных областях // Информационные и математические технологии в науке и управлении. Иркутск: ИСЭМ СО РАН. 2016. № 4–1.— С. 27–36.

2. Turban E. Decision Support and Expert Systems: Management Support Systems. 4th Edition, Macmillan Publishing Company, New York (1995).

3. Hitzler P., Krötzsch V., Rudolph S. Foundations of Semantic Web Technologies. Chapman & Hall/CRC, 2009. 455 p.

4. Загорулько Ю. А., Загорулько Г. Б., Боровикова О. И. Технология создания тематических интеллектуальных научных интернет-ресурсов, базирующаяся на онтологии // Программная инженерия. 2016. № 2. С. 51–60.

5. Загорулько Г. Б., Массель Л. В. Разработка интеллектуальной СППР по предотвращению угроз энергетической безопасности //Вестник СибГУТИ. 2019. № 3. С. 70–79.

ТЕХНОЛОГИЯ ПРОГНОЗА УРОЖАЙНОСТИ КАРТОФЕЛЯ

С ИСПОЛЬЗОВАНИЕМ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ

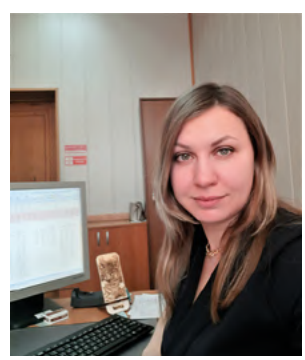
Сибирский научно-исследовательский гидрометеорологический институт,
630099, г. Новосибирск,
ул. Советская, 30

Новосибирский государственный аграрный университет,
630039, г. Новосибирск,
ул. Добролюбова, 160

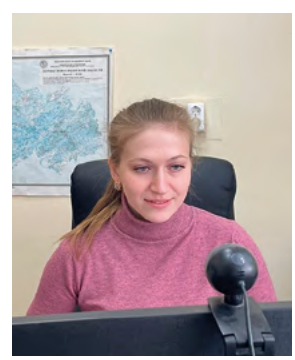
Пищимко Олеся Игоревна
старший научный сотрудник
тел.: +7-952-926-68-79, pishchimko@sibnigmi.ru

Сибирский федеральный научный центр агробиотехнологий Российской академии наук,
630501, Новосибирская область, р. п. Краснообск,
ул. Центральная, Президиум

Гарафутдинова Людмила Вячеславовна
младший научный сотрудник
тел.: +7-952-926-69-93, lv.garafutdinova@mail.ru



Работа в лаборатории. Пищимко Олеся Игоревна



Работа в лаборатории. Гарафутдинова Людмила Вячеславовна

Прогнозирование урожайности сельскохозяйственных культур — это сложная и важная задача, зависящая от природных и агротехнологических факторов, и при этом необходимая для эффективного использования природных ресурсов. Следовательно, благодаря заблаговременному прогнозированию урожая принимаются решения по своевременному принятию мер по повышению урожайности сельскохозяйственных культур (применению комплекса агротехнологических мероприятий), от которых зависит продовольственная безопасность страны.

Была разработана технология расчета урожайности картофеля, в основе которой лежат модели множественной регрессии. В качестве предикторов применялись стандартные агрометеорологические данные наземной сети станций (осадки, дефицит насыщения воздуха, температура), данные дистанционного зондирования Земли (спутниковые снимки) и средняя областная урожайность картофеля по сельскохозяйственным предприятиям, взятая из статистических сборников субъектов региона. Работа и ее апробация проводилась по сельхозпредприятиям Иркутской области. Из агрометеорологических ежегодников была проведена выборка данных, в которую вошли 20 репрезентативно расположенных метеостанций наземной сети (Качуг, Киренск, Залари, Усть-Илимск, Новожилкино, Худоеланское, Железногорск, Тангуй, Тай-

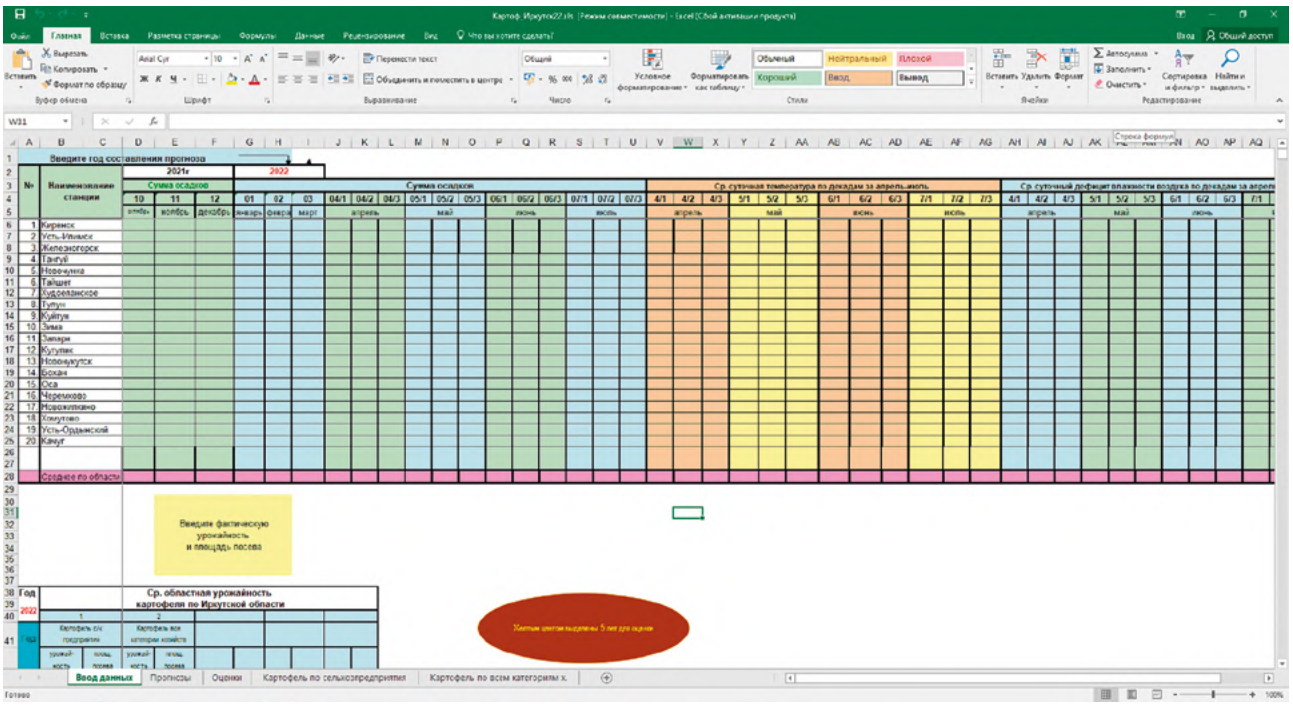


Рис. 1. Страница вводных данных.

шет, Тулун, Куйтун, Кутулик, Новонукутск, Бохан, Зима, Оса, Черемхово, Новочунка, Хомутово и Усть-Ордынский). Для разработки моделей использовались многолетние ряды данных с 1981 по 2020 г. и 13-летний ряд подекадных значений NDVI. На следующем этапе составлялись базы данных и рассчитывались корреляционные матрицы в программе Microsoft Excel, которые отражают зависимость метеоданных, NDVI и урожайности исследуемой культуры.

Прогноз урожайности картофеля в представленной технологии составляется в стандартный принятый в Росгидромете срок 1–3 августа. При построении моделей прогноза урожайности картофеля выбор временного суммирования NDVI и агрометеорологических параметров определяли в зависимости от фазы вегетации растений. Параметры моделей множественной регрессии определяли с помощью метода наименьших квадратов в среде программирования MATLAB.

Оценка моделей проведена согласно руководящему документу «Методические указания по проведению производственных (оперативных) испытаний новых и усовершенствованных методов гидрометеорологических и гелиогеофизических прогнозов» РД 52.27.284–91. Модели с использованием индекса NDVI показали достаточно высокую оправдываемость, которая составила более 90% (в среднем за три года авторских испытаний).

Полученные модели вошли в автоматизированную технологию расчета урожайности картофеля, созданную в программе Microsoft Excel, которая позволяет ежегодно производить расчет прогноза урожайности картофеля. Пользователь вносит в вводную форму агрометеорологические параметры, и расчет происходит автоматически, выдавая в результате прогноз урожайности, прогноз валового сбора и оценку методического прогноза согласно РД.

Вовремя предоставленный прогноз урожайности картофеля позволяет руководству региона принять своевременные решения и уменьшить экономические потери, снизить риски природного характера, обеспечить продовольственную безопасность. С учетом интенсификации сельскохозяйственной отрасли и изменений климата требуется агрометеорологическое обеспечение, соответствующее запросам сельхозпроизводителей и руководящих органов региональной власти.

Предложенная технология может быть использована специалистами сельскохозяйственного производства (агрономами, экономистами и др.).

МОДЕЛИРОВАНИЕ РАСПРОСТРАНЕНИЯ ЭПИДЕМИЙ

С УЧЕТОМ ЭКОНОМИЧЕСКОЙ И СОЦИАЛЬНОЙ СИТУАЦИИ В РЕГИОНЕ

Федеральное государственное бюджетное учреждение науки Институт математики им. С. Л. Соболева Сибирского отделения Российской академии наук (ИМ СО РАН), 630090, г. Новосибирск, пр. Академика Коптюга, 4

Федеральное государственное бюджетное учреждение науки Институт вычислительной математики и математической геофизики Сибирского отделения Российской академии наук (ИВМиМГ СО РАН) 630090, г. Новосибирск, пр. Академика Лаврентьева, 6

Криворотько Ольга Игоревна
кандидат физико-математических наук, старший научный сотрудник, начальник научно-исследовательского отдела Международного математического центра ИМ СО РАН
тел.: +7-983-303-30-83, krivorotko.olya@mail.ru

Кабанихин Сергей Игоревич
доктор физико-математических наук, член-корреспондент РАН, директор Международного математического центра ИМ СО РАН
тел.: +7-913-910-34-78, ksi52@mail.ru

Зятьков Николай Юрьевич
кандидат технических наук, научный сотрудник, начальник инновационного отдела Международного математического центра ИМ СО РАН
тел.: +7-953-873-33-12, n.y.zyatkov@math.nsc.ru

Неверов Андрей Вячеславович
инженер, Институт вычислительной математики и математической геофизики СО РАН
тел.: +7-913-798-25-19, a.v.neverov@g.nsu.ru

На основе обработки статистических данных методами машинного обучения и применения математического аппарата можно оценить сценарии распространения инфекционного заболевания в регионе и эффективность сдерживающих его распространение мер на месяц вперед по неполным и неточным статистическим данным.

Разработан комплекс программ ЭПИДЕМИЯ для моделирования и построения сценариев распространения COVID-19 в регионе на 45 дней с учетом социально-экономических процессов (вакцинации, введенных ограничений, открытых данных соцсетей), в котором учитываются:

- распределение населения в исследуемом регионе на возрастные группы с шагом по 10 лет;
- четыре вида контактов: домохозяйства, образовательные учреждения, рабочие и общественные места;
- ежедневные данные о выявленных умерших, протестированных, госпитализированных, индекс самоизоляции от Яндекс; эпидемиологические оценки ВОЗ: средняя продолжительность пребывания в инкубационном периоде, инфицирования, пребывания в стационаре интенсивной терапии;

Рис. 1. Научная группа на 2-й конференции математических центров международного уровня, МИАН, Москва, ноябрь 2022 (слева направо): Сурнин П.С. (студент НГУ), Шишленин М.А. (профессор РАН), Криворотько О.И. (к.ф.-м.н.), Кабанихин С.И. (чл.-корр. РАН), Неверов А.В. (студент НГУ).

Рис. 2. Научная группа на рабочем семинаре по распространению COVID-19 в Красноярске, сентябрь 2021 (слева направо): Зятьков Н.Ю. (к.т.н.), Шайдуров В.В. (чл.-корр. РАН), Кабанихин С.И. (чл.-корр. РАН).

- введенные в регионе ограничения и послабления (масочный режим, вакцинация, карантин), появление новых штаммов заболевания и их характеристик.

Комплекс ЭПИДЕМИЯ определяет:

- параметр заразности, количество бессимптомных носителей вируса, шанс быть протестированным в день t;
- значение числа репродукции (среднее количество человек, зараженное инфицированным больным) в течение всего периода моделирования;
- изменение количества выявленных инфицированных, госпитализированных, подключенных к аппарату ИВЛ, умерших индивидуумов в каждый день t.

оценивает:

- сценарии изменения параметра заразности в зависимости от социально-экономических процессов, появления новых штаммов;
- ожидаемое количество выявленных случаев заболевания (положительный тест ПЦР), общее число инфицированных и пациентов в реанимации с 90% доверительным интервалом в зависимости от различных вве-

Рис. 3. Сценарии распространения COVID-19 для новых выявленных случаев в Новосибирской области с 07.08 по 12.09.2022: базовый (синяя линия), с частичной изоляцией рабочих мест (оранжевая линия), повышение мобильности граждан в общественных местах с 21.08 (зеленая линия), появление нового штамма (красная линия). Реальные данные для валидации моделирования обозначены красными точками.

денных мер (закрытие общественных мест, переход на удаленную работу, уменьшение количества контактов в учебных заведениях и общественных местах и другие).

Результат работы программы

Построены сценарии изменения количества выявленных случаев заболевания COVID-19 в Новосибирской области с 07.08.2022 по 12.09.2022:

- базовый сценарий на момент моделирования 06.08.2022 (не учитываются введенные ограничения и возможные новые штаммы вируса);
- частичный 30% перевод рабочих мест на удаленную работу с 21.08;
- увеличение мобильности граждан в общественных местах с 21.08 в преддверии школьных ярмарок;
- учет появления нового штамма с большей контагиозностью (Кентавр).

В рамках модели количество выявленных случаев в Новосибирской области на 30.08.2022 ожидалось по сценарию появления нового штамма в регионе (Кентавр) — 700 случаев, реальные данные — 671 выявленный случай (ошибка прогнозирования составляет 4%).

Средний порядок ошибки вычисления: 5–20% в зависимости от текущей эпидемиологической ситуации.

Также возможно проанализировать влияние различных вариантов вируса (например, искусственно ввести в популяцию несколько агентов с новым штаммов вируса в определенный день).

Постепенно будет введен функционал по анализу вакцинации и изменению иммунитета, однако это пока недоработанные расширения, так как человечество в целом еще не до конца уверено в том, насколько быстро исчезает иммунитет, насколько вакцина защищает от повторного заражения и насколько уменьшает вероятность осложнений.

Представленный программный комплекс решает проблему оценки эффективности воздействия социально-экономических мер на скорость распространения эпидемии и загруженности медицинской системы. Новизна научного результата состоит в построении эффективных новых алгоритмов численного решения возникающих обратных задач по неполным и неточным данным эпидемиологических и социально-экономических процессов, сочетающих в себе методы машинного обучения и искусственного интеллекта, природоподобные алгоритмы и баесовские методы. На данный комплекс программ получены 5 свиде-

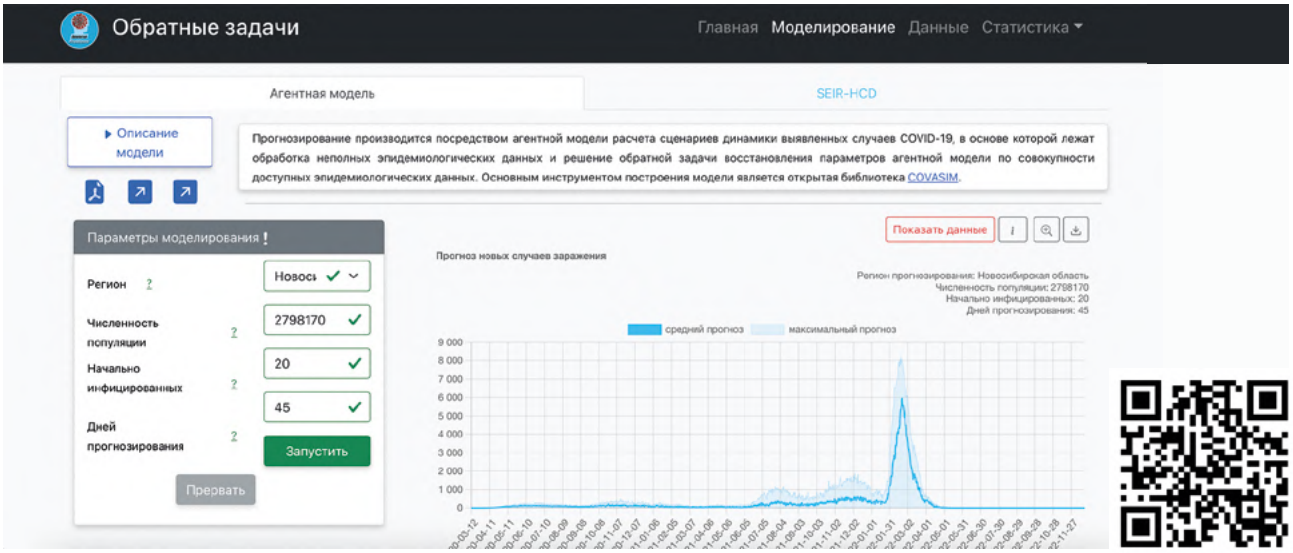


Рис. 4. Веб-страница моделирования распространения эпидемии COVID-19 в Новосибирской области, разработанного научной группой сайта. На представленном графике изображен базовый прогноз распространения COVID-19 в Новосибирской области до конца 2022 года.

тельств о регистрации программ для ЭВМ в системе Роспатент.

Для описания процессов распространения эпидемий были развиты два принципиально разных подхода к построению математических моделей:

- компартментный (моделирование «сверху вниз»), основанный на законе действующих масс и описываемый системой дифференциальных уравнений, где вся популяция разделяется на группы со схожими признаками, связанными между собой параметрами перехода;
- агентный (моделирование «снизу вверх»), исследующий поведение децентрализованных агентов и то, как такое поведение определяет поведение всей системы в целом. Включает в себя клеточные автоматы, элементы теории игр, сложных систем, мультиагентных систем и эволюционного программирования, методы Монте-Карло.

Также наряду с классическими подходами к построению моделей в период пандемии COVID-19 были разработаны модели машинного обучения, основу которых составляют нейронные сети, регрессионные модели и модели фильтрации.

Математические модели характеризуются своими параметрами, описывающими особенности распространения эпидемии в исследуемом регионе. Основная задача построения более точных сценариев распространения эпидемии состоит в уточнении параметров моделей по дополнительной информации об эпидемиологическом и связанных социально-экономических процессах.

Основные отечественные и зарубежные аналоги математических моделей описания распространения эпидемии COVID-19:

Базельский университет, Швейцария. Реализуется структурированная по возрастам компартментная модель распространения коронавирусной инфекции, основанная на 9 дифференциальных уравнениях с возможностью варьирования параметров модели.

ИВТ СО РАН, Новосибирск. Моделируется распространение COVID-19 в Москве, Новосибирской области, Германии, Франции и Италии, используются расширенная камерная и агентная модели, параметры которых идентифицируются на основе опубликованной статистики. Строятся прогнозы не только числа регистрируемых заболевших, вылечившихся и умерших, но и количества свободных койко-мест, аппаратов искусственной вентиляции легких (ИВЛ) и других характеристик,

Криворотько Ольга Игоревна

Личный комментарий от разработчика:

«Математический аппарат и методы искусственного интеллекта способны на основе неполных и неточных статистических данных о ходе распространения эпидемии и социально-экономических процессов спрогнозировать возможную вспышку эпидемии и ее нагрузку на систему здравоохранения и экономику региона».

необходимых для эффективного управления ситуацией в условиях эпидемии.

Богота, Колумбия. Создана на основе агентной модели с учетом географических особенностей города (расположение школ, медицинских учреждений, общественных мест) и социальной дистанции между агентами с возможностью варьирования параметров.

Хельсинки, Финляндия. На основе агентной модели учитывается различный возраст агентов и семь возможных состояний прогрессирования заболевания, а также рассматривается случайная структура контактов, т.е. агенты взаимодействуют между собой хаотично.

США. На основе агентной модели со случайной структурой с возможностью идентификации параметров для конкретного региона. Данный программный комплекс послужил основой создания собственного программного комплекса.

Предполагаемый интерес для внедрения

Разработка может быть интересна научным сотрудникам для тестирования и оценки алгоритмов прогнозирования распространения эпидемий с учетом социально-экономических процессов, а также региональным и муниципальным органам власти для мониторинга эпидемиологической ситуации (например, мэрии города Новосибирска).

Рис. 5. Совещание научной группы отделения математических наук РАН по вопросам оценки прогноза распространения COVID-19 с учетом социальных, экономических и экологических процессов в кабинете заместителя директора по науке ИМ СО РАН (слева направо): Кабанихин С.И., Криворотько О.И., Шананин А.А., Шайдуров В.В.

КЛЕТОЧНО-АВТОМАТНАЯ МОДЕЛЬ ДИНАМИЧЕСКИХ ПРОЦЕССОВ
В ГАЗОУГОЛЬНОМ РАСТВОРЕ

СибГИУ
Первый вуз Кузбасса

Федеральное государственное бюджетное образовательное учреждение высшего образования «Сибирский государственный индустриальный университет»
654007, г. Новокузнецк, Центральный р-н, ул. Кирова, 42

Немцев Александр Юльевич
преподаватель
тел.: +7-960-923-10-05, alexg02r02l@gmail.com

Калашников Сергей Николаевич
доктор технических наук, доцент, профессор
тел.: +7-913-422-25-50, s.n.kalashnikov@yandex.ru

Сеченов Павел Александрович
кандидат технических наук, доцент
тел.: +7-950-582-07-05, sechenov.p.a.1989@gmail.com

Калугин Константин Павлович
студент
тел.: +7-913-295-26-37, kpkalugin@yandex.ru

Современный Центр цифровых компетенций СибГИУ.

Защита проекта Немцева А.Ю. в финале программы «Я в деле» на площадке СибГИУ.

Более точное прогнозирование аварий при подземных работах.

Программа предназначена для моделирования массоообмена в конечном объеме угля, а именно диффузии и равновесной сорбции. Использована 2D-модель клеточного автомата с окрестностью Марголуса со способом обработки крайних клеток — обрабатываются реализуемый через дополнительный как внешний, так и внутренний контур и модифицированными функциями перехода с по часовой и против часовой на по вертикали и по горизонтали. Программа представляет собой набор макросов для работы в среде MS Excel, пользовательский интерфейс и средства, необходимые для моделирования вышеуказанного процесса.

Разработанный алгоритм решения уравнения применим для исследования других физических процессов: тепло- и электропроводности, магнитных и электромагнитных потоков и др. В таких компьютерных экспериментах наблюдается самоорганизация материи, и эта само-

организация вытекает из локальных взаимодействий, правила которых задаются, а структура возникает сама.

Исполнение как двух тактов в одном такте, так и диффузии, и конвекции в одной не последовательной, а вероятностной суперпозиции упрощает запись и возможна к применению в других клеточных автоматах.

Использование дополнительных краевых клеток для реализации внешних краевых условий и способа обработки краевых клеток упрощает их замену, так как нужно заменить лишь их, вместо замены переходных функций краевых клеток, что удобнее при масштабировании и быстрее.

Программа может использоваться для прогнозирования движения газа в угле на угледобывающих предприятиях, в том числе аварии, например, газоугольный выброс.

Основные характеристики

1. Среднее количество шагов клеточного автомата, за которое моделируется равномерное распределение концентрации газа по конечному объему угля — 100 шагов.

2. Среднее время расчета — 1 минута.

3. Объем программы для ЭВМ: 15 700 байт.

4. Количество модулей — 3.

5. Среда разработки — MS Excel.

6. Язык программирования — VBA (Visual Basic for Application).

Свидетельство о регистрации программы для ЭВМ «Клеточно-автоматная модель динамических процессов в газоугольном растворе», Немцев А. Ю., Калашников С. Н., Сеченов П. А., Калугин К. П., № 2023614385, 01.03.2023.

Результаты проведенных численных экспериментов показали, что клеточные автоматы обладают рядом возможностей с точки зрения качественного моделирования газодинамических процессов. Разработанная методология компьютерного моделирования эволюции состояний конечного объема угля предусматривает использование клеточного автомата как инструмента для исследования.

Оценка уровня готовности технологии (TRL)

TRL3 — проведён вычислительный эксперимент.

Переход на TRL4 путем модификации и верификации модели с помощью лабораторного эксперимента, для чего необходимо разработать лабораторный стенд. Стоимость разработки лабораторного стенда — 500 000 р.

Преимущества и уникальность

Уменьшение количества аварий за счет увеличение точности прогнозирования вероятности аварий на угольных шахтах.

Новизна

1. Разработана математическая модель движения газа сквозь газоугольный раствор из пор и внешней среды, в которой учитывается неоднородность среды, что позволяет прогнозировать динамику протекания динамических процессов.

2. Разработан численный метод и алгоритм расчета поля концентрации газа по конечному объему угля, отличающийся учетом проникновения газа из внешней среды и газовой фазы в твердую фазу дисперсной системы и позволяющий спрогнозировать газоугольный выброс.

3. Разработан комплекс программ, реализующий разработанный алгоритм вычисления движения газа в газоугольном растворе, отличающийся визуализацией результатов расчёта и позволяющий вычислять с высокой вычислительной скоростью поле концентрации газа в конечном объеме угля.

Контекст

Клеточно-автоматное моделирование — представление объекта исследования в виде набора клеток, каждая клетка которого описывает состояние в конечном объеме объекта исследования и изменения состояния клетки по правилам перехода. Дискретность системы позволяет естественно распараллелить вычисления модели и представить объекты, имеющие прерывистую природу — например пористый материал. Теория клеточных автоматов была разработана в 1940-х годах Вольфрамом и Тьюрингом и доработана фон Нейманом, Муром, Маргулосом и Тоффоли. Технологический

прогресс позволил применять на практике дискретные модели, такие как клеточные автоматы и нейронные сети.

Клеточные автоматы обладают таким преимуществом над нейронными сетями, как интерпретируемость. Во многих случаях необходимо учитывать внутренние свойства материала, среди которых главные — топология среды и характер взаимодействия стенок пор с проходящим через них газом или жидкостью. А рассмотрение пористых материалов как сплошных сред, характеризуемых определенным коэффициентом пористости, не удовлетворяет исследователей горных пород.

Иногда при решении довольно простых дифференциальных уравнений возникают проблемы, связанные со сложными граничными условиями. Так, при решении параболического уравнения, описывающего диффузию в области со сложной границей (например, в пористой среде) возникают серьезные трудности. Поскольку вычислительные мощности современных суперкомпьютеров позволяют решить эту задачу, то становятся актуальными новые методы моделирования, такие как мелкозернистый параллелизм, а конкретнее — клеточные автоматы.

В связи с появлением компьютеров с высокой степенью параллельности для моделирования диффузионных процессов широко используются КА, при помощи которых были решены некоторые проблемы газодинамики.

Ближайшие аналоги

Исследование и моделирование дисперсных систем, например конечного объёма угля, затруднено при использовании для описания этих систем дифференциальных уравнений, пригодных для описания непрерывных однородных сред. Это обусловлено необходимостью задания краевых условий (нелинейностей) для неоднородных сред (дисперсных систем), то есть при нарушении сплошности в материале, что можно адекватно реализовать с помощью клеточных автоматов.

Ряд авторов занимаются данной проблематикой.

1. Клеточно-автоматные модели газоугольного раствора, разработанные Макаровым Павлом Васильевичем и его учениками в Томском государственном университете, описывают физико-кинетические процессы, протекающие в растворе (например, разрушение материалов), а не физико-химические процессы.

2. Анатолий Васильевич Шестопалов остановился на начальном этапе разработки клеточно-автоматных моделей газоугольного раствора, то есть на разработке математических моделей, но не реализуя их как компьютерные модели.

3. Бандман Ольга Леонидовна применяет клеточные автоматы в области физико-химических процессов, но не в конкретно газоугольных растворах, а вообще в дисперсных системах.

4. Ряд других авторов останавливаются только на физических процессах, не рассматривая химические. Сейчас в практике применения клеточно-автоматных методов доминирует метод подвижных клеточных автоматов для моделирования деформации и разрушения материалов, но не химических превращений.

Существуют большое количество моделей искусственных наноматериалов (дисперсные системы), но мало природных. Из моделей природных наноматериалов основная часть сконцентрирована на физико-кинетических процессах (деформация), моделируемых с помощью метода подвижных клеточных автоматов. Однако такие модели не учитывают физико-химические процессы (процессы пространственной динамики), моделируемые с помощью метода блочно-синхронных клеточных автоматов. Из существующих же моделей пространственной динамики нет ни одной, которая была бы реализацией на компьютере такой дисперсной системы, как газоугольный раствор.

Имеющиеся же не клеточно-автоматные модели углей плохо учитывают нарушения сплошности (нелинейности) дисперсных систем в связи с описанием этих систем дифференциальными уравнениями: это вызвано трудностями при необходимости учёта краевых условий, что связано дисперсностью моделируемой среды. Проблема учета нарушения сплошности снимается при использовании для моделирования дисперсных систем клеточных автоматов. Кроме этого существенным отличием клеточных автоматов является их естественная параллельность, что дает преимущество при расчётах на распределённых системах.

Немцев Александр Юльевич

Личный комментарий от разработчика:

«У нас выше точность прогнозирования за счет применения системы, лишенной таких недостатков, как ошибка дискретизации и неинтерпретируемости».

В связи с этим далее будут приведены лишь качественные характеристики разрабатываемой компьютерной модели, отличающие её от ближайших аналогов. Такими характеристиками являются:

- новый стохастический блочно-синхронный метод разбиения клеточного массива, заключающийся в обмене значениями клеток в блоке не по или против часовой стрелки, как с использованием окрестности Марголуса, а по горизонтали и вертикали, что позволяет разделить блок на логические полублоки для более плотной загрузки вычислительной мощности отдельных ядер в матричном процессоре (СБИС, нейроускоритель, графический ускоритель) на один такт;
- новый метод (внешний контур) задания внешних краевых условий, работающий только в паре с ранее изложенным новым методом стохастического блочно-синхронного клеточного автомата, заключающийся в четырех режимах работы, два из которых являются режимами обработки краевых клеток:
- краевые клетки либо не обрабатываются, либо обрабатываются, а два других режима — это симуляции либо притока, либо оттока энергии, материи, информации; новый метод задания внутренних краевых условий с помощью композиции двух клеточных автоматов над одним клеточным массивом, где первый автомат задаёт топологию моделируемого объекта, а второй автомат имитирует течение эволюции внутренних состояний моделируемого объекта.

Рис. 1. Окрестность Мура, фон Неймана, эквивалентная окрестности Марголуса, Марголуса в однотактном исполнении.

Предполагаемый интерес для внедрения

Проектным институтам, внедряющим на угольные шахты системы мониторинга без системы прогнозирования или с меньшей точностью.

Отрасли для потенциального применения: все виды подземных работ.

Чего ждем от индустриального партнера

Натурные данные с их системы мониторинга, устройство этой системы, доля совмещения наших систем, партнерское соглашение и дорожную карту по данным работам.

На рис. 1 изображены: окрестность Мура; фон Неймана; эквивалентная окрестности Марголуса; эквивалентная окрестности Марголуса без отступа идентична окрестности фон Неймана; Марголуса в однотактном исполнении, выделенные пунктирной линией, клетка, вместе со своими соседними клетками создающая окрестность, выделена черным цветом, а для окрестности Марголуса в однотактном исполнении является пересечением двух блоков из двухтактного исполнения.

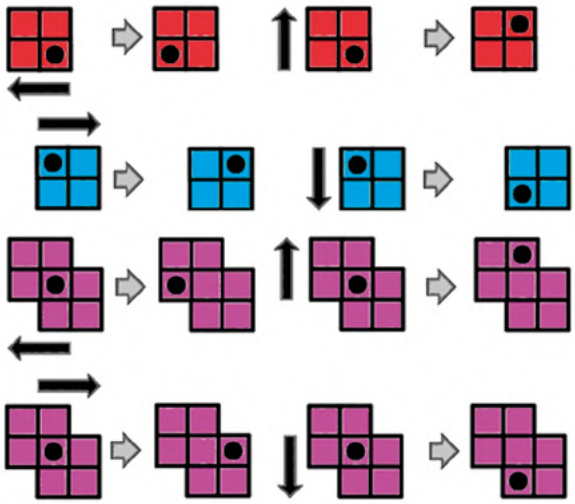


Рис. 2. Изменение состояния клетки на пересечении двух блоков по и против часовой стрелки в верхнем или нижнем блоке.

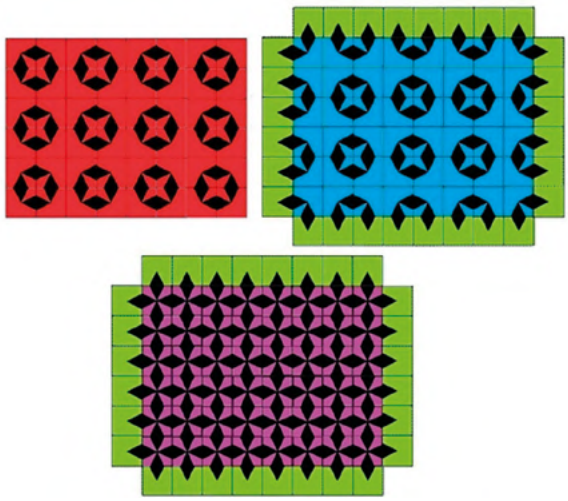


Рис. 3. Схема разбиения массива клеток КА с окрестностью Марголуса в двухтактном исполнении, с обработкой крайних клеток — замыкание и однотактном исполнении, с обработкой крайних клеток — замыкание.

На рис. 2 проиллюстрировано изменение состояний в двухтактном обмен происходит равновероятный по часовой и против часовой стрелки, внутри блока и однотактном исполнении равновероятно обмен происходит или в верхнем, или в нижнем блоке с пересечение в одной клетке, равновероятный по часовой и против часовой стрелки, внутри блока.

На рис. 3 изображена схема разбиения массива клеток КА с окрестностью Марголуса в двухтактовом исполнении, со способом обработки крайних клеток – замыкание, и однотактное исполнение со способом обработки крайних клеток – замыкание. Замыкание реализуется не изменением функции перехода крайних клеток, а дополнительным контуром из клеток конфигурация которого зависит от окрестности КА, с помощью этого контура можно задать внешние краевые условия и обработку краевых клеток: приток энергии контур равен 1; отток энергии контур равен 0; не обрабатывается (изолирован) равен граничащими с контуром клеток; замыкание (тор) однородная среда равен граничащими с контуром клеток с обратной стороны, контур выделен контурной линией.

На рис. 4 показаны результаты серии вычислительных экспериментов компьютерной модели динамических процессов в газоугольном растворе с 4 микропорами. Использована 2D-модель КА с окрестностью Марголуса в однотактном исполнении, со способом обработки крайних клеток – замыкание. Размеры всего массива 16x16 клеток, один массив со «свободным газом» 4x4. На границах «твердое тело – газ» применялось условие (3). Клетки с толстыми черными границами «твердое тело», клетки с тонкими границами «твердое тело – газ», клетки с бледно-серыми границами «свободный газ». Серые клетки – наличие молекулы газа (заполненные), белые клетки – отсутствие молекулы газа (незаполненные).

На рис. 5 показаны результаты серии вычислительных экспериментов компьютерной модели динамических процессов в газоугольном растворе с 1 микропорой. Использована 2D-модель КА с окрестностью Марголуса, со способом обработки крайних клеток – не обрабатываются. Размеры всего массива 29x19 клеток, один массив со «свободным газом» 13x7. На границах «твердое тело – газ» применялось условие (3). Клетки с толстыми черными границами «твердое тело», клетки с тонкими границами «твердое тело – газ», клетки с бледно-серыми границами «свободный газ». Серые клетки – наличие молекулы газа (заполненные), белые клетки – отсутствие молекулы газа (незаполненные).

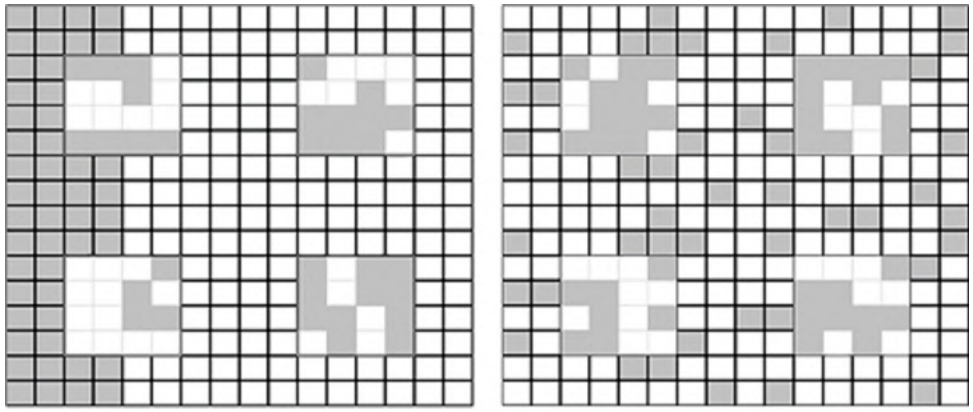


Рис. 4. Результаты вычислительного эксперимента: начальное состояние; сотый шаг.

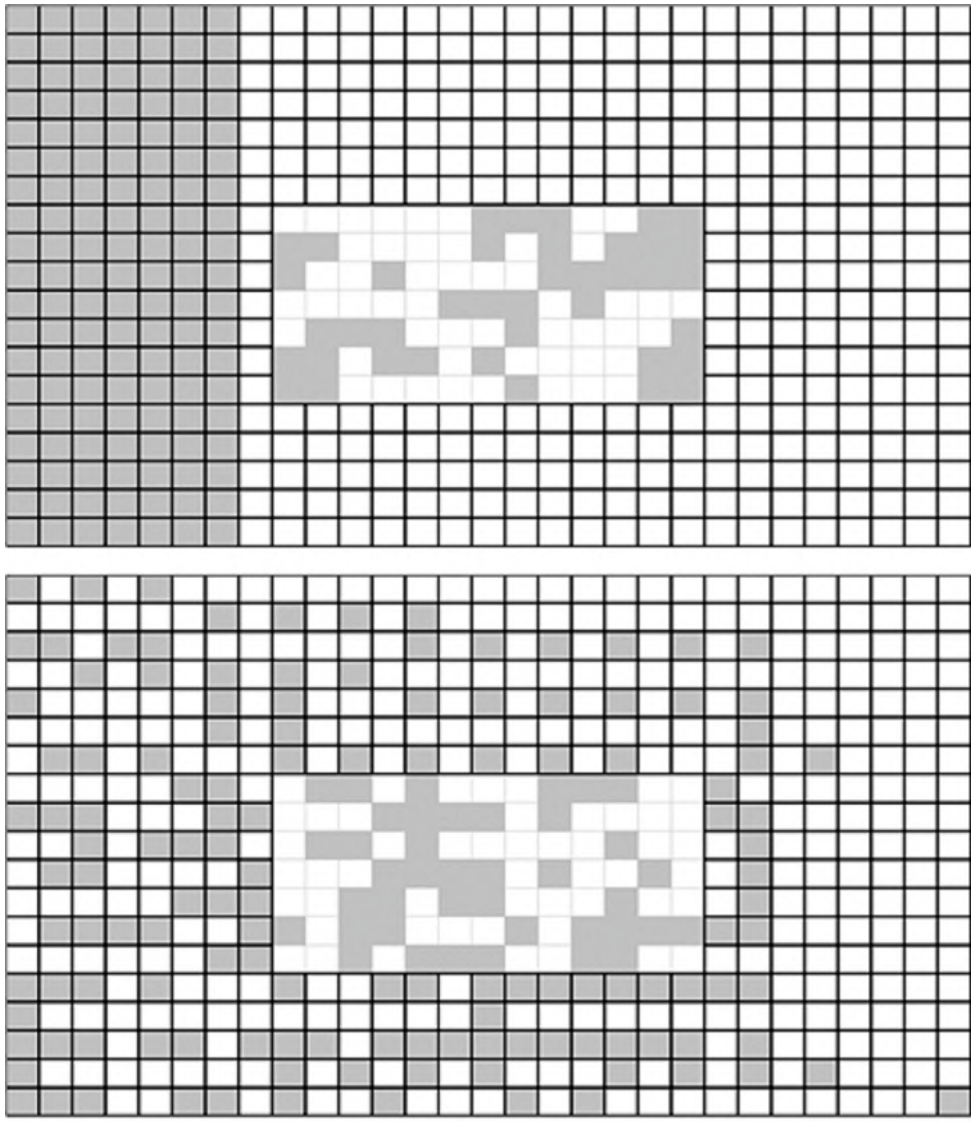


Рис. 5. Результаты вычислительного эксперимента: начальное состояние; сотый шаг

ОПТИЧЕСКИЕ СИСТЕМЫ ПАНОРАМНОЙ ДИАГНОСТИКИ ПОТОКОВ «ПОЛИС»

Общество с ограниченной ответственностью «Сигма-Про»
630090, г. Новосибирск,
ул. Инженерная, 28

Федеральное государственное бюджетное учреждение науки Институт теплофизики им. С. С. Кутателадзе Сибирского отделения Российской академии наук
630090, г. Новосибирск,
пр. Академика Лаврентьева, 1

Амосов Константин Александрович
директор
тел.: (383) 373-21-92, info@polis-instuments.ru

Скрябиков Сергей Владимирович
коммерческий директор
тел.: (383) 373-21-97, comdir@polis-instuments.ru

Бильский Артур Валерьевич
заместитель директора, кандидат физико-математических наук
тел.: +7-913-905-89-18, bilsky@itp.nsc.ru

Гобызов Олег Алексеевич
младший научный сотрудник
тел.: (383) 316-50-38, oleg.a.g.post@gmail.com

Амосов Константин Александрович, директор

Бильский Артур Валерьевич, заместитель директора, кандидат физико-математических наук

Скрябиков Сергей Владимирович, коммерческий директор

Гобызов Олег Алексеевич, младший научный сотрудник

Компания «Сигма-Про» [1] является ведущим российским производителем специализированных оптических систем «ПОЛИС», предназначенных для проведения измерений в потоках жидкости и газа. Измеряемые параметры потоков: поле скорости, поле температуры, поле давления, размеры и количество капель/пузырей, формы объектов, поле деформации объектов и другие величины. Основными пользователями систем «ПОЛИС» являются научные организации, отраслевые институты и конструкторские бюро, такие как МГУ, НГУ, ИТ СО РАН, ИТПМ СО РАН, ЦНИИ им. Крылова, НИТИ имени Александрова, ОКБ «ГИДРОПРЕСС», РФЯЦ ВНИИТФ, ЦНИИ машиностроения, ЦАГИ им. Н. Е. Жуковского, АО «ОДК-Авиадвигатель» и другие.

Основой для развития линейки продуктов компании многие годы служат совместные разработки с Институтом теплофизики им. С. С. Кутателадзе СО РАН (ИТ СО РАН). В системах «ПОЛИС» используются так называемые панорамные оптические методы, иногда называемые методами «количественной визуализации». Для реализации этих методов обычно применяется комплекс оборудования, состоящий из средств регистрации изображений — специализированных научных или промышленных видеокамер, средств освещения потока - лазеров нескольких типов или диодных осветителей, а также компьютера, устройства синхронизации и другого вспомогательного оборудования. Вычисление характеристик течения происходит на основе регистрируемых изображений при помощи специального программного обеспечения ActualFlow.

Первым продуктом линейки «ПОЛИС», представленным в 2007 году, были системы анемометрии по изображениям частиц (Particle Image Velocimetry, PIV) [2]. Метод PIV позволяет буквально «увидеть» движение жидкости, используя микрочастицы, добавляемые в поток. Процесс измерения заключается в том, что частицы добавляются в несущую фазу (жидкость или газ), затем поток освещается лазерным излучением, а на камеру регистрируются изображения подсвеченных частиц. Имея последовательность таких снимков, с помощью специального программного обеспечения можно проанализировать перемещение частиц, вычислить скорость и направление движения жидкости или газа на разных участках изображения. Если первые системы предназначались для измерений в потоках жидкости, то позднее были разработаны конфигурации для измерений в потоках газа

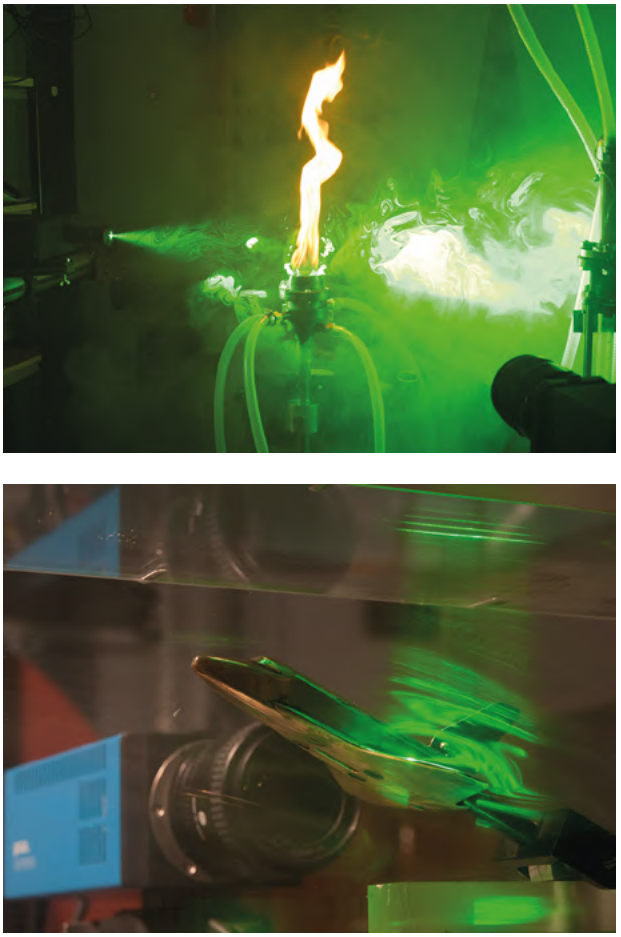


Рис. 1. PIV-измерения в потоке газа и в пламени.

и реагирующих течениях (высокотемпературном газе). Также были реализованы варианты стереоскопических PIV-систем (позволяют получать значения трёх компонент скорости в сечении потока), томографических PIV-систем, которые позволяют получать значения трёх компонент скорости в объеме потока, и скоростных систем с частотой регистрации до десятков тысяч измерений в секунду. Несмотря на расширение областей применения систем «ПОЛИС», PIV-системы остаются одним из основных продуктов компании.

Линейка систем «ПОЛИС» на сегодняшний день

Со временем всё большее число методов, как использующихся в ИТ СО РАН для собственных исследовательских работ, так и разработанных под нужды пользователей, включалось в продуктовую линейку систем «ПОЛИС». На сегодняшний день, помимо PIV, существуют модифи-

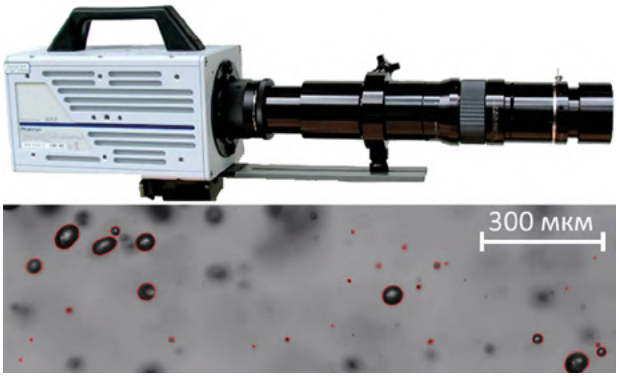


Рис. 2. Высокоскоростная камера с объективом-микроскопом и образы капель на снимке.

кации систем для измерения полей температуры (температурная лазерно-индуцированная флуоресценция, ЛИФ) [3], исследования процессов смешения в течениях жидкости и газа (концентрационный ЛИФ) [4], диагностики двухфазных пузырьковых и капельных течений [5], исследования полей давления на поверхности моделей в задачах аэродинамики (метод люминесцентных преобразователей давления, PSP) [6, 7], и ряд других. Для реализации методов разрабатывались специализированные оптомеханические устройства, совершенствовалась аппаратная база систем, и главное — развивалось программное обеспечение для автоматизации системы и обработки полученных данных.

В последние годы пополнился арсенал методов, предназначенных для диагностики газокапельных течений и процессов распыла жидкостей форсунками. В этой области существует потребность в определении размеров капель, их скорости, а также различных интегральных характеристик, например геометрии факела (конуса) распыла. Для определения размера и скорости капель в системах «ПОЛИС» используется метод теневой фотографии с большим оптическим увеличением. Специальная оптика и устройства подсветки позволяют получать изображения с увеличением порядка 10X, то есть непосредственно видеть капли и различать их форму [8]. Наглядность метода является его важным конкурентным преимуществом.

Теневая фотография с меньшим увеличением, охватывающая всю область факела распыла, позволяет определить его угол раскрытия, наклон, а в случае импульсного впрыска — визуализировать развитие факела и определить скорость его фронта.

В определении характеристик распыла находит своё место и метод лазерно-индуцированной флуоресценции. При добавлении специального красителя — люминофора в рабочую жидкость метод позволяет получать распределение концентрации жидкости в сечении факела распыла и более детально исследовать его структуру.

Программное обеспечение систем «ПОЛИС»

Ключевую роль в работе систем играет программное обеспечение, необходимое для управления оборудованием и обработки полученных изображений. В постоянном сотрудничестве с ИТ СО РАН «Сигма-Про» поддерживает и развивает программный пакет ActualFlow [9], являющийся главной составной частью систем «ПОЛИС». ActualFlow обеспечивает настройку и управление аппаратными средствами, сбор и хранение данных, широкий спектр инструментов обработки и визуализации.

Наиболее важными в ActualFlow являются модули обработки данных. Такие модули строятся на основе опыта проведения исследований, интегрируют в себе наборы алгоритмов для реализации каждого из методов и обеспечивают последовательный процесс измерения от регистрации исходных данных до получения целевых физических величин. Так, например, модуль ЛИФ включает в себя процедуры пространственной калибровки изображений, калибровки яркости от температуры или концентрации, ряд алгоритмов предобработки изображений и непосредственно процедуры вычисления полей температуры или концентрации на основе построенных калибровок.

В 2022 году пакет ActualFlow получил очередное масштабное обновление — пользовательский интерфейс стал более интуитивным и современным, а логика работы с данными и их представления была адаптирована для работы с результатами высокоскоростных измерений, всё чаще востребованных пользователями.

На сегодняшний день в России работает более 80 измерительных систем «ПОЛИС». Благодаря постоянному развитию, качественному научному и техническому сопровождению и сотрудничеству с научными организациями в ходе внедрения новых решений комплексы «ПОЛИС» компании «Сигма-Про» стали успешным про-

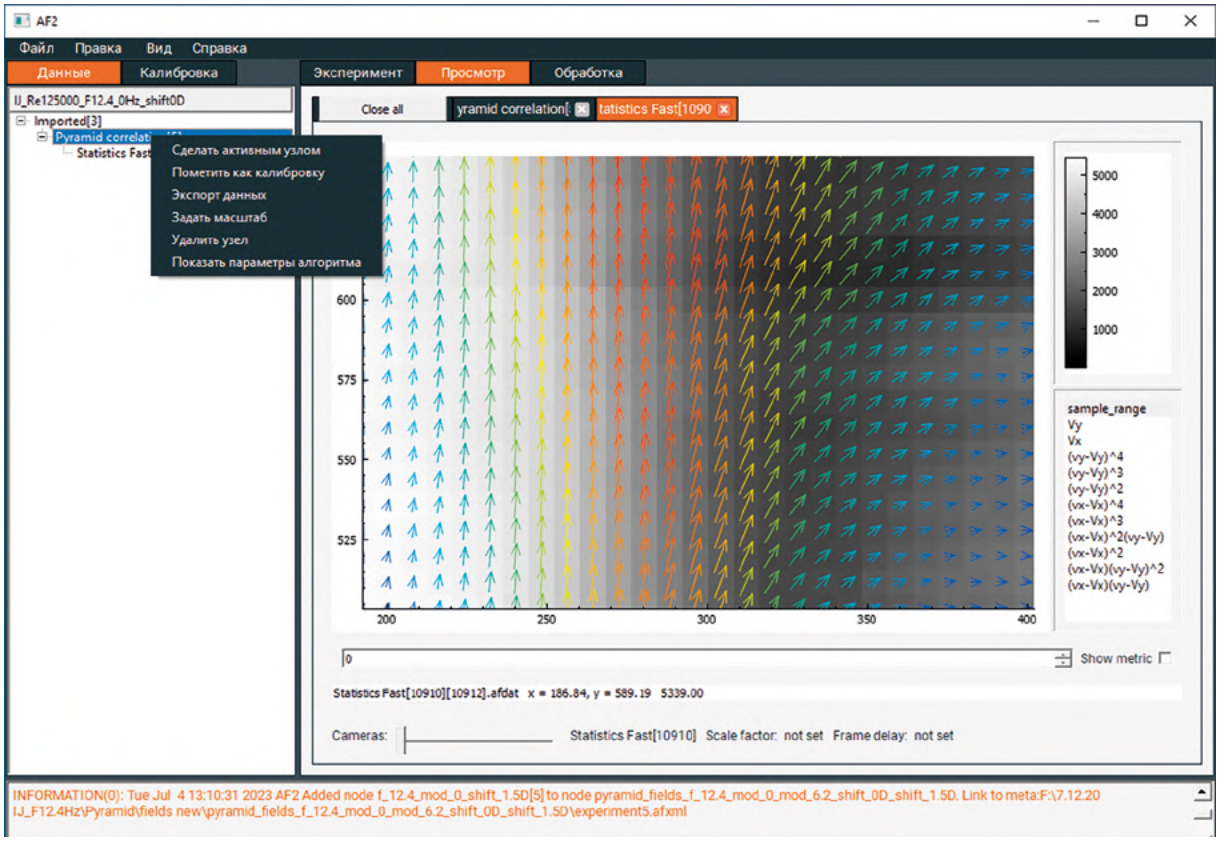


Рис. 3. Основное окно пользовательского интерфейса ПО ActualFlow.

дуктом, который используется не только в академических исследованиях, но и для решения задач стратегического сектора промышленности, позволяя глубже понять и совершенствовать технологические процессы.

Литература

1. http://polis-instruments.ru

2. Raffel M., Willert C. E., Scarano F., Kähler C., Wereley S. T., Kompenhans J. 2018. Particle image velocimetry, a practical guide 3rd ed // Springer Int. Publishing. 669 p.

3. Nebuchinov A.S., Lozhkin Y. A., Bilsky A. V., Markovich D. M., Combination of PIV and PLIF methods to study convective heat transfer in an impinging jet, Experimental Thermal and Fluid Science. Vol. 80, P. 139–146.

4. Чикишев Л.М., Дулин В. М., Гобызов О. А., Лобасов А. С., Маркович Д. М. (2017). Исследование смесеобразования в модели камеры сгорания ГТУ с использованием панорамных оптических методов. Теплофизика и аэромеханика. Т. 24, № 3, С. 357–364.

5. Alekseenko S.V., Dulin V. M., Markovich D. M., Pervunin K. S. (2015). Experimental investigation of turbulence modification in bubbly axisymmetric jets. Journal of Engineering Thermophysics. Vol. 24, No. 2. P. 101–112.

6. Liu T. Pressure and temperature sensitive paints / T. Liu, J. P. Sullivan. Berlin: Springer, 2005.

7. Budovsky A.D., Sidorenko A., Luzgin N.K., Gobyzov O. A., Ryabov M.N., Korotko M.S. (2022). Application of the PSP technique to study the structure of transonic flow in rectangular channel., International Conference on the Methods of Aerophysical Research, Book of Abstracts, P. 34–35.

8. Gobyzov O, Ryabov M, Bilsky A. Study of Deformation and Breakup of Submillimeter Droplets' Spray in a Supersonic Nozzle Flow (2020) Applied Sciences. Vol. 10, № 18, P. 6149.

9. Е.К. Ахметбеков, А. В. Бильский, Ю. А. Ложкин, Д. М. Маркович, М. П. Токарев, А. Н. Тюрюшкин (2006). Система управления экспериментом и обработки данных, полученных методами цифровой трассерной визуализации (ActualFlow). Вычислительные методы и программирование. Т. 7, С. 79–85.

ОПТИКО-ИНФОРМАЦИОННАЯ ДИАГНОСТИКА ДЛЯ ЭНЕРГЕТИКИ, ПРОМЫШЛЕННОСТИ И ТРАНСПОРТА

Федеральное государственное бюджетное учреждение науки Институт теплофизики им. С. С. Кутателадзе Сибирского отделения Российской академии наук
630090, г. Новосибирск, пр. Академика Лаврентьева, 1

Меледин Владимир Генриевич
доктор технических наук, профессор, главный научный сотрудник, лауреат Государственной премии РФ, премии Правительства РФ
тел.: (383) 330-87-82, meledin@itp.nsc.ru, director@ioit.ru

Одним из важнейших производственных факторов современной экономики, наряду с ресурсами, стала диагностическая информация. Знание точных параметров технологических процессов позволяет технологиям оптимально работать в отрицательном поле допусков и выполнять обслуживание техники и транспорта по фактическому состоянию. Результат — существенная экономия энергии, средств и материалов, уменьшение количества отходов и улучшение эффективности, повышение безопасности, значительный экономический и социальный эффект в рамках отдельных предприятий и отраслей.

Оптико-информационная диагностика, позволяющая реализовать обозначенные преимущества, принципиально мультидисциплинарна в отличие от IT. Оптико-информационная диагностика особо перспективна для науки, техники и инновационных промышленных технологий, а также для верификации расчетных кодов и компьютерных моделей в энергетике, промышленности и на транспорте. Современные оптико-информационные диагностические системы, не уступающие лучшим зарубежным аналогам, простые и безопасные в использовании, крайне необходимы для отечественного научного и инженерного образования мирового уровня и подготовки кадров инновационной экономики России.

Предложены принципы оптико-информационной диагностики, ставшие фундаментальной научной базой широкого класса созданных систем. Разработаны методы формирования, приема и программно-аппаратной обработки сложных оптико-электронных и акустооптических сигналов и изображений в условиях фотонного ограничения, способы их обработки, в том числе на основе вэйвлетов и мультифракталов, методы динамического мониторинга трехмер-

Рис. 1. Разработка и создание средств прецизионной лазерной диагностики утвержденного первичного государственного эталона единицы скорости воздушного потока России ГЭТ-150-2012.

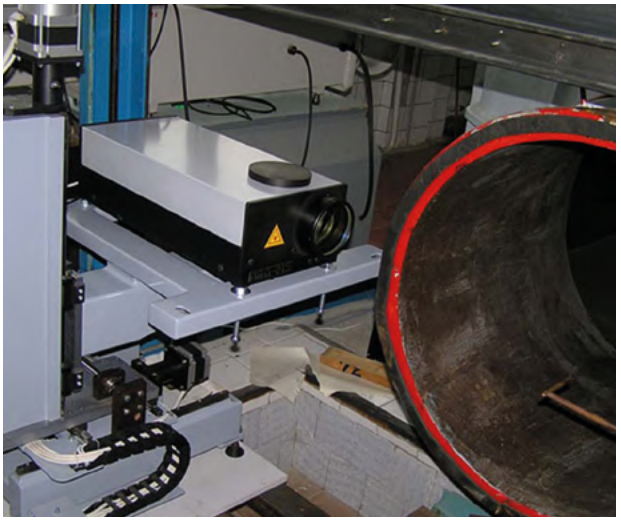


Рис. 2. Государственный эталон ГЭТ-150-2012 с ЛАД-015.

ной геометрии объектов. Впервые созданы мощные полупроводниковые стабилизированные лазерные излучатели для доплеровской анемометрии и научного приборостроения (совместно с ЛФТИ им. А. Ф. Иоффе). Создана линейка современных диагностических систем для регистрации полей физических величин в газожидкостных многофазных потоках «ЛАД-0**» на основе новых анаморфотных оптических схем и современных мощных полупроводниковых лазеров с ограниченной пространственной и временной когерентностью. Прототип этих систем опередил появление лучшего зарубежного аналога (DANTEC) на три года.

Системы внедрены на многих предприятиях и в научно-образовательных учреждениях, используются в рамках приоритетных направлений развития науки, технологий и техники в Российской Федерации. Они задействованы в энергетике для изучения кавитационных и нестационарных процессов и снижения эрозии в элементах гидроагрегатов, исследования обтекания и оптимизации формы лопаток паровых турбин (ОАО «Силовые машины — ЛМЗ»), оптимизации процессов в топочных камерах, горелках (ИТ СО РАН), на транспорте для натурных испытаний авиационных и ракетных двигателей (ФГУП ВО «Внештехника», СГАУ), испыта-

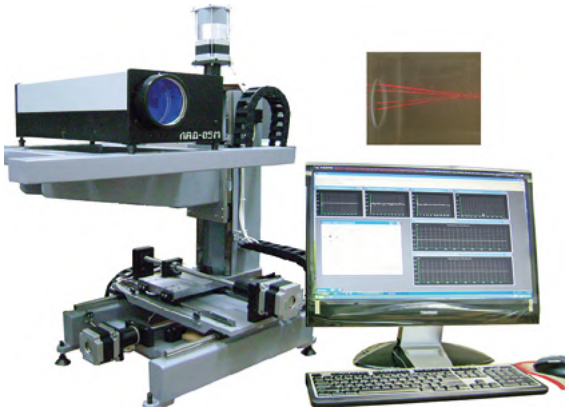


Рис. 3. Лазерные полупроводниковые доплеровские ИИС (LDSA, 2001-2011), импортозамещение.

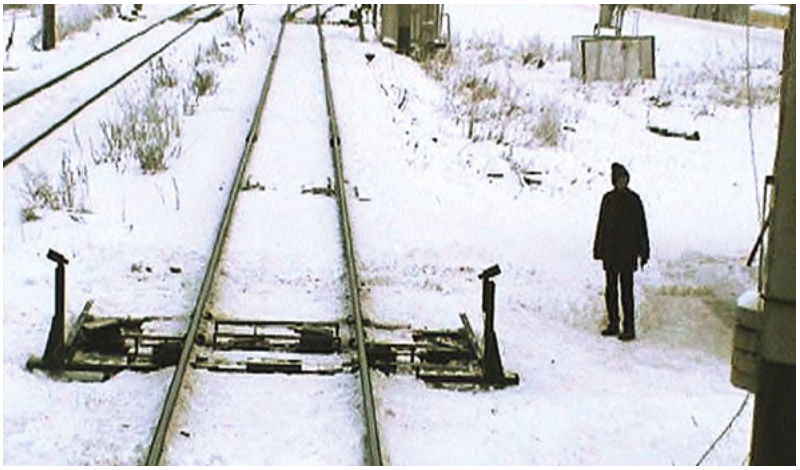


Рис. 4. Информатика транспорта. Лазерный диагностический комплекс для измерения колесных пар вагонов на подходах к станции. Натурные испытания. Срок разработки (НИР+ОКР) - 2 месяца. Опытный образец на Транссибирской магистрали. ПТО «Агрыз», ноябрь 2001 г.

ний моделей летательных аппаратов и изделий ракетно-космической техники (КГТУ-КАИ, ИТПМ СО РАН и др.), оптимизации и диагностики топливных форсунок авиационных двигателей, процессов перемешивания в камерах сгорания перспективных авиационных двигателей (ОАО «Авиадвигатель»), оптимизации формы надводных и подводных судов (ФГУП «Крыловский государственный научный центр»), диагностики распространения лесных пожаров (ТПУ, ФГБУ ВНИИПО МЧС России) и т.д.

Создан ряд лазерных оптико-информационных диагностических систем для учета и технологического контроля проката на металлургических предприятиях, обеспечивших получение в реальном времени адекватной и ранее отсутствовавшей информации о скорости и геометрии проката, позволившей автоматизировать учет, улучшить стабильность, безопасность и экономичность технологий, обеспечить выход продукции повышенного качества. Новейшая система для металлургии «Лазерный радиационно-безопасный измерительный комплекс для мониторинга геометрических параметров горячего и холодного проката ЛАД-0Р 3» предназначена для бесконтактных, точных и безопасных для персонала измерений толщины движущегося горячего листового проката в режиме реального времени, осуществляемых в тяжелых условиях окружающей среды. Наукоемкая диагностика реализована на самой современной элементной базе фотоники и электроники. Измерение толщины проката производится на основе синхронной дифференциальной лазерной облачной триангуляции с многоуровневой спектральной и пространственной фильтрацией. Применены инновационные патентно чистые научные и технические решения, оригинальные алгоритмы собственной разработки, обеспечившие погрешность измерений на уровне 10^{-6} в лабораторных условиях и 10^{-5} в реальных условиях горячего цеха металлургического завода.

Созданные оптико-информационные диагностические системы внедрены и успешно работают на крупнейших промышленных предприятиях черной металлургии России (НТМК (1993 — настоящее время), ЗСМК (1992–2001), НМЗ (1986 — настоящее время), Первоуральский новотрубный завод (2011— настоящее время)). Первая отечественная лазерная диагностическая система для мониторинга параметров горячего проката «Квазар-М» непрерывно эксплуатировалась 20 лет на Новосибирском металлургическом заводе, метрологически обеспечив выпуск более 160 миллионов 8-метровых стальных труб. Затраты на внедрение технологии оптико-информационной лазерной диагностики горячего проката на крупнейшем предприятии отрасли — Нижнетагильском металлургическом комбинате — окупились за несколько суток использования, а предложенные и реализованные оптико-информационные системы и технологии непрерывно и успешно работают в металлургическом производстве России почти 30 лет. По отзывам российских металлургов, по ряду важнейших параметров (рабочее расстояние до проката более 3 000 мм, глубина зоны точных измерений более 400 мм,

Рис. 5. Лазерный «Комплекс» после снегопада и сдача автором «Комплекса» комиссии министерства, ПТО «Агрыз», 2001 год.

общая надежность), как на момент внедрения, так и в настоящее время, созданные и внедренные оптико-информационные комплексы демонстрируют ряд существенных преимуществ перед лучшими мировыми аналогами (Джордж Келк — Канада, TSI–США и др.).

Крупным достижением в сфере практического использования наукоемкой оптико-информационной диагностики на транспорте является впервые предложенный и реализованный в 2001 году способ мониторинга геометрии железнодорожных колесных пар проходящих поездов на основе синхронной лазерной регистрации мгновенных радиус-векторов до поверхности колесных пар и сверхбыстродействующей лазерной триангуляции. Автором и его коллективом в совместной лаборатории ИТ-КТИ НП впервые создан действующий прототип системы «Комплекс», обеспечивший бесконтактный, дистанционный, в темпе движущегося поезда и в условиях реальных перепадов погодных условий, температур, пыли, грязи и прочих помех надежный информационный мониторинг геометрии железнодорожных колес (ПТО «Агрыз» Горьковской железной дороги, 2001 г.). На этой базе ООО «ЦТТ» при сопровождении КТИ НП СО РАН производятся и широко внедряются серийные всепогодные оптико-информационные лазерные системы «Комплекс» для автоматического бесконтактного контроля колесных пар грузовых вагонов при движении поезда со скоростью до 60 км/ч. Они успешно эксплуатируются на всех железных дорогах России и зарубежья в качестве штатного средства контроля геометрии колёсных пар, в России с 2002 года внедрено около ста систем. Предотвращая аварии и обеспечивая безопасность перевозок, диагностические системы «Комплекс» контролируют геометрию колёсных пар 60 млн вагонов в год, причем отцепляется для поверки и ремонта до 40 тыс. вагонов в год. Подтвержденный экономический эффект от внедрения систем в ОАО «РЖД» к 2015 году составил более 10 млрд руб.

В области фундаментальной метрологии и обеспечения единства средств измерений успешно выполнена разработка и созданы оптико-информационные средства прецизионной лазерной диагностики первичного Государственного специального эталона единицы скорости воздушного потока России ГЭТ-150–2012 (ВНИИМ, г. Санкт-Петербург) на основе методов прецизионной полупроводниковой лазерной доплеровской анемометрии. Уровень выполненных работ подтвержден 11-месячными международными ключевыми сличениями национальных эталонов единицы скорости воздушного потока CIPM CCM-FF-K3, в которых принимали участие первичные эталоны США, Японии, России, Кореи, Сингапура и Тайваня. Результаты международных сличений доказали, что обновленный Государственный первичный эталон России вышел на лидирующие позиции в мире.

Благодаря уникальному сочетанию мультидисциплинарного кадрового потенциала, достижений фундаментальной науки, доступности

6

8

9

10

7

11

специализированных конструкторских бюро и экспериментального производства в новосибирском Академгородке проведены разработки, выпуск конструкторской документации и освоение малых серий целого ряда оптико-информационных диагностических систем для энергетики, промышленности и транспорта. Реализация и внедрение выполнялись на основе частно-государственного партнерства ИТ СО РАН, академических институтов и ряда организаций, в том числе возглавляемой автором специализированной внедренческой организации ОАО «Институт оптико-электронных информационных технологий» (ОАО «ИОИТ»), лауреатом главной национальной бизнес-премии Торгово-промышленной Палаты РФ «Золотой Меркурий» в области инновационной деятельности. Организация ведет ориентированные НИОКР на стыке лазерных, нано-, био-, электронных и информационных технологий, доводя результаты работ до промышленных внедрений, решая проблему наукоемкого импортозамещения и обеспечивая существенный рост эффективности и безопасности многих инновационных промышленных и оборонных технологий. ОАО «ИОИТ» внесено в реестры Федерального справочника «Оборонно-промышленный комплекс России» (ФСОПК 6/94, 9/27, 11/30, 16/156) как предприятие, имеющее важное значение в развитии и укреплении оборонно-промышленного потенциала РФ.

Научные и практические результаты работ по оптико-информационной диагностике для энергетики, промышленности и транспорта неоднократно отмечались многими российскими и международными почетными наградами. В их числе: Государственная премия РФ за выдающиеся достижения в области науки и технологий; премия Правительства РФ в области науки и техники; премия имени А. Н. Косыгина в области науки, техники и организации производства; победа в Конкурсе русских инноваций медиахолдинга «Эксперт» в номинации «Лучшая промышленная инновация России»; золотые медали Московских международных салонов инноваций и инвестиций; золотые медали и дипломы международных промышленных выставок «Сибполитех»; медаль Innovations for investments to the future Американо-российского делового союза (ARBU) в рамках международной программы Golden Galaxy — за развитие инновационной де-

ятельности, техническое перевооружение производства и активное участие в процессе внедрения инновационных технологий, а также множество других коллективных и индивидуальных наград.

Результаты широкого применения разработанных наукоемких оптико-информационных диагностических систем и технологий и отзывы пользователей подтверждают их высокий технический уровень, не уступающий, а по ряду параметров (точность, функциональность, требования к квалификации персонала, научно-техническое сопровождение) —превосходящий лучшие мировые аналоги. Созданные импортозамещающие системы защищены 75 российскими и международными патентами, успешно внедрены и используются в различных отраслях экономики, соответствуя мировому уровню исследований и разработок и глобальной конкурентоспособности Российской Федерации на направлениях, определенных национальными научно-технологическими приоритетами для энергетики, промышленности и транспорта.

Литература

1. Меледин В. Г. Информационная оптоэлектронная диагностика. Наука и инновационные промышленные технологии / отв. редактор С. В. Алексеенко. — Новосибирск: Академиздат, 2015. — 173 с.

2. Меледин В. Г. Информатика оптоэлектронных измерений: наука и инновационные промышленные технологии // Новосибирск: Изд-во ИТ СО РАН, 2008. — 75 с.

3. Меледин В.Г. 3D-диагностика — просто, точно, доступно / Наука из первых рук, 2010. — № 2. — С. 23–24.

4. V. Meledin. Optoelectronic Measurements in Science and Innovative Industrial Technologies // Optoelectronic Devices and Properties. ISBN 978-953-307-204-3. — Chapter 18. — Suffolk, UK: INTECH, 2011. — P. 373–399.

5. Strelnik A.S., Dvoinishnikov S. V., Meledin V. G., Kabardin I. K. (2022) Development of a Doppler anemometry method for diagnosing two-phase flows in a liquid metal medium. In: Sergiyenko O. (eds) Optoelectronic Devices in Robotic Systems. pp. 167–194 Springer, Cham, Switzerland

6. Dvoynishnikov S.V., Kabardin I. K., Meledin V. G. Advanced Phase Triangulation Methods for 3D Shape Measurements in Scientific and Industrial Applications. In: Sergiyenko O., Flores-Fuentes W., Mercorelli P. (eds) Machine Vision and Navigation. Springer, 2020.

7. S.V.Dvoynishnikov, V. G. Meledin. Optoelectronic Differential Cloudy Triangulation Method for Measuring Geometry of Hot Moving Objects / Optoelectronics in Machine Vision-Based Theories and Applications, M.Rivas-Lopez, O.Sergiyenko, W.Flores-Fuentes, J. C. Rodríguez-Quiñonez, August 2018, p. 49–78.

8. M.A.Bolshukhin, A. V. Budnikov, E. I. Shmelev, D. A. Kulikov, A. V. Loginov, N. A. Pribaturin, P. D. Lobanov, V. G. Meledin, A. S. Suvorov, A. V. Stulenkov. Dynamic measurements of the flow and structure oscillations to validate FSI calculations // Nuclear Engineering and Design, Volume 381, September 2021, 111336.

9. Алексенко С.В., Меледин В. Г., Двойнишников С. В., Аникин Ю. А., Бакакин Г. В., Главный В. Г., Кабардин И. К., Климов А. В., Кротов С. В., Куликов Д. В., Наумов И. В., Павлов В. А., Прибатурин Н. А., Рахманов В. В., Садбаков О. Ю., Чубов А. С., Щепихин И. В. Развитие оптико-информационных методов, систем и технологий бесконтактной диагностики гидродинамических процессов и трехмерных пространственных структур // Вихревые явления и их влияние на процессы переноса / под ред. С. В. Алексеенко и И. В. Наумова; Ин-т теплофизики им. С. С. Кутателадзе СО РАН. — Новосибирск: ИПЦ НГУ, 2018. — 388 с. — С. 8–70.

10. V. Meledin. Informatics of Optoelectronic Measurements: Science and Innovative Industrial Technologies // Journal of Engineering Thermophysics, 2009. — Vol.18. — No. 2. — P. 99–128.

11. Dvoynishnikov S.V., Rakhmanov V. V., Kabardin I. K., Meledin V. G. Phase triangulation method with spatial modulation frequency optimization // Measurement, Vol.145 (2019). — P. 63–70.

Рис. 6. Награда победителя конкурса «Научный прибор года – 2011», Москва

Рис. 7. X МСИИ – 2 медали аверс

Рис. 8. Награда лауреата конкурса русских инноваций - статуэтка Надежда.

Рис. 9. Медаль ARBU 2009 – 2012 год.

Рис. 10. Медаль Санкт-Петербург 2011 – 2012 год.

Рис. 11. Медаль национальной премии «Меркурий», 2 стороны, 2012 год.

2

Искусственный интеллект и большие данные

64	Цифровая библиотека буддийского наследия: методы машинного обучения в изучении древних книжных памятников
70	Набор данных EMNIST TCyrillic для тренировки моделей машинного обучения распознавания символов тюркских письменностей, построенных на основе расширенной кириллицы
73	Нейросетевая система диагностики острого инсульта на основе компьютерного анализа КТ-изображений головного мозга
75	Интеллектуальная технология прогнозирования инсоляции
78	Технология для персонализированной диагностики нейроонкологических заболеваний с помощью искусственного интеллекта
82	Метод оценки эффективности применения геофизических исследований для решения геологоразведочных задач, основанный на применении искусственного интеллекта и построении трехмерной геолого-геофизической модели
85	Проблемы обеспечения правовой защиты персональных данных при использовании технологии искусственного интеллекта в медицинской деятельности

ЦИФРОВАЯ БИБЛИОТЕКА БУДДИЙСКОГО НАСЛЕДИЯ:

МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ В ИЗУЧЕНИИ ДРЕВНИХ КНИЖНЫХ ПАМЯТНИКОВ

Федеральное государственное бюджетное учреждение науки Институт монголоведения, буддологии и тибетологии Сибирского отделения Российской академии наук (ИМБТ СО РАН)
670047, г. Улан-Удэ,
ул. Сахъяновой, 6

Федеральное государственное автономное образовательное учреждение высшего образования «Новосибирский национальный исследовательский государственный университет»
630090, г. Новосибирск,
ул. Пирогова, 2

Ринчинов Олег Сергеевич
доктор исторических наук, координатор Центра восточных рукописей и ксилографов
тел.: (3012) 43-02-15, orinch2@mail.ru

Окунев Алексей Григорьевич
кандидат химических наук, директор Высшего колледжа информатики НГУ
тел.: (383) 373-11-99, okunev73@mail.ru

Рис. 1. Здание Центра восточных рукописей и ксилографов ИМБТ СО РАН

Обширная цифровая библиотека, отражающая обширность и многообразие буддийской письменной культуры, в которой представлены многие тысячи произведений разных жанров, созданная с применением наиболее актуальных информационных технологий, включая методы искусственного интеллекта.

Целью проекта является раскрытие научного и социального потенциала буддийской культуры на основе полномасштабной цифровизации и комплексного мультидисциплинарного исследования с применением методов искусственного интеллекта памятников письменности на тибетском и старописьменном монгольском языках, представленных в уникальных коллекциях Центра восточных рукописей и ксилографов Института монголоведения, буддологии и тибетологии Сибирского отделения Российской академии наук. В результате выполнения проекта будет создана обширная цифровая библиотека, отражающая обшир-

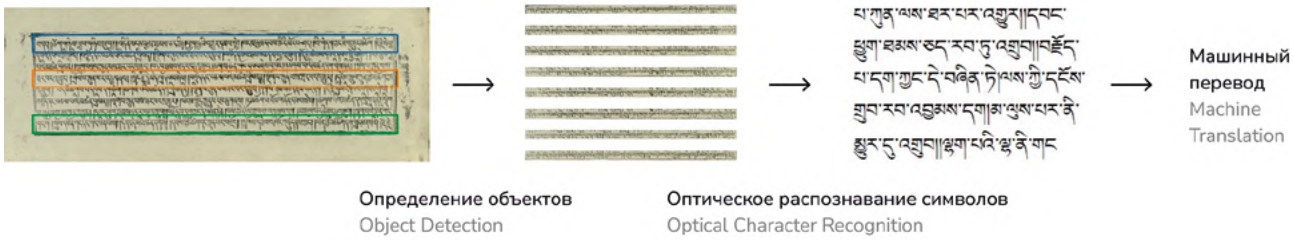


Рис. 2. Задачи искусственного интеллекта в проекте «Цифровая библиотека буддийского наследия: методы машинного обучения в изучении древних книжных памятников».

ность и многообразие буддийской письменной культуры, в которой найдут место многие тысячи произведений разных жанров.

Научная значимость и актуальность изучения письменного наследия тибетского буддизма определяются огромным влиянием буддийской культуры в странах Восточной и Центральной Азии. С одной стороны, буддийская культура имеет важное структурообразующее значение для социально-политической жизни в азиатских обществах, обеспечивая непрерывность передачи духовных знаний и традиций, сохранение культурной и национальной идентичности. В то же время благодаря универсальным идеям и ценностям она способствовала сложению общего культурного пространства, становлению новых направлений межкультурного взаимодействия, технологических и торгово-экономических обменов, включая такие яркие исторические явления, как Шелковый путь, Чайный путь и др. В связи с неуклонным ростом значения этого макрорегиона в современном глобальном мире освоение достижений буддийской культуры представляется важным и актуальным.

Россия имеет длительный опыт проведения взвешенной политики межкультурного и межконфессионального взаимодействия во многом благодаря вхождению в ее состав народов, исповедующих буддизм и имеющих многосторонние связи со странами буддийского ареала. Памятники буддийского наследия, в том числе обширные книжные коллекции, являются национальным достоянием нашей страны и прямым мостом духовного и культурного взаимодействия с азиатскими соседями. В современных условиях использование этого потенциала предоставляет России существенные преимущества при формировании стратегии «мягкой силы» в важнейшем макрорегионе Восточной и Центральной Азии, учитывающем региональный историко-культурный контекст. Поэтому обращение к памятникам буддийского наследия, представленных в российских коллекциях, на современном академическом и технологическом уровне с упором на активное применение цифровых технологий позволит найти решение многих задач социокультурного развития России, укрепит историческую роль и место нашей страны в культурном пространстве Азии.

Изучение памятников письменности, хранящихся в фондах ИМБТ СО РАН, с применением методов искусственного интеллекта и цифровых инструментов анализа текстов актуально в свете важных для современного поликультурного общества проблем освоения культурного наследия народов буддийской Азии, включающих изучение философских, этических и естественнонаучных концепций, которые получили отражение в обширном корпусе буддийской книжности и традиционной медицине.

В основе проекта лежит междисциплинарный подход, в котором особый упор сделан на применение современных цифровых и когнитивных технологий, что позволит представить наследие буддийской цивилизации на принципиально новом для российской науки уровне. Создание цифровой библиотеки и ввод в научный оборот уникального книжного собрания ИМБТ СО РАН станут научным достижением мирового значения.

Важнейшим преимуществом проекта, обуславливающим его уникальный характер не только в российской, но и международной практике, является применение актуальных ныне методов «глубокого обучения» искусственного интеллекта, которые основаны на использовании

программируемых нейронных сетей, имеющих несколько последовательных слоев преобразований и самообучения. Эти методы позволяют эффективно решать задачи декодирования сканированных текстов, машинного перевода, содержательного анализа текстов.

В ходе реализации проекта модели глубокого (машинного) обучения будут разработаны для решения следующих задач:

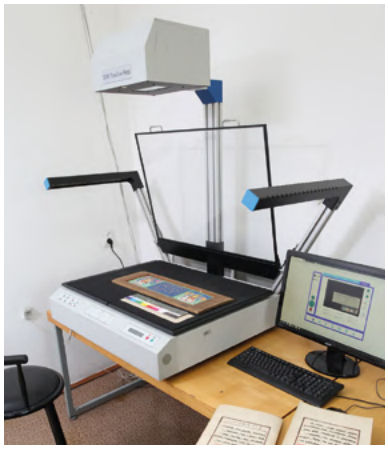
- поиск/детектирование областей с текстом на изображении;
- распознавание найденного изображения с текстом в машиночитаемые текстовые данные (декодирование);
- машинный перевод распознанного текста на русский.

Для разработки моделей детектирования ОИ на оцифрованных тибетских и монгольских рукописях и ксилографах будут созданы в значительном объеме (до нескольких тысяч) обучающие материалы — датасеты, включающие изображение с текстом, графическую разметку (прямоугольники с текстом для обучения моделей типа SSD либо полигоны для обучения моделей типа MaskRCNN) и текстовые аннотации. Для разработки моделей машинного перевода должны быть сформированы объемные параллельные корпусы текстов (тибетско-русский и монгольско-русский).

В результате выполнения проекта будут разработаны передовые средства цифровизации письменных памятников на основе нейросетевых методов «глубокого обучения». искусственного интеллекта — такие как системы оптического распознавания символов тибетской и монгольской письменностей и машинного перевода с тибетского и старописьменного монгольского языков. Запланировано выполнение поточной дешифровки оцифрованных рукописей и ксилографов из фондов ИМБТ СО РАН. На этой основе будет сформирована обширная цифровая библиотека буддийских писаний и введен в научный оборот значительный цифровой корпус произведений буддийской классики, снабженный современными инструментами анализа текстов. В его составе войдут в научный оборот ценные письменные источники на тибетском и старописьменном монгольском языках, включая уникальные разновременные издания буддийского канона, что позволит раскрыть научный, культурный и социальный потенциал буддийской цивилизации.

Особое внимание уделяется следующим рукописным и ксилографическим источникам: а) тибетоязычные и монголоязычные источники по традиционной медицине; б) сочинения бурятских и монгольских авторов на старописьменном монгольском и тибетском языках; в) книги, изданные монастырскими печатными дворами Забайкалья. Будет сформирован многоязычный параллельный корпус сочинений по традиционной медицине. Результаты проекта реализуются в размещенных в сети Интернет общедоступных веб-сервисах, предоставляющих доступ к оцифрованным памятникам письменности, электронным тезаурусам и специализированным базам данных, инструментам анализа текстов.

Коллектив проекта, состоящий из сотрудников ИМБТ СО РАН и НГУ, при поддержке Сибирского отделения РАН реализовал пилотный проект «Подготовка датасета для обучения baseline моделей оптического распознавания символов тибетского языка», который позволил впервые применить технологию «глубокого обучения» искусственного интеллекта (deep learning) для создания работающего прототипа системы оптического распознавания символов тибетской письменности. На основе отобранных материалов из числа сканированных раритетных ксилографических изданий на тибетском языке были подготовлены 500 обучающих датасетов, включающих размеченные изображения книжных страниц и текстовые аннотации, выполненные специалистами-тибетологами ИМБТ СО РАН с использованием веб-приложения LabelStudio, развернутого на сервере НГУ. Датасеты применялись для обучения двухстадийной R-CNN OCR-модели. Полученный в результате машинного обучения алгоритм достиг 94% точности распознавания символов. В результате реализации пилотного проекта заложена методическая база для создания полнофункциональной системы оптического распознавания символов тибетской письменности, на основе которой осуществляется поточное декодирование сканированного текста на тибетском языке. Была показана перспективность применения технологий искусственного интеллекта для изучения рукописей и ксилографов на восточных языках. Результаты проекта вошли в доклад Президента РАН академика А. М. Сергеева, представленный Президенту Российской Федерации В. В. Путину в марте 2022 года и обсуждались на их личной встрече.



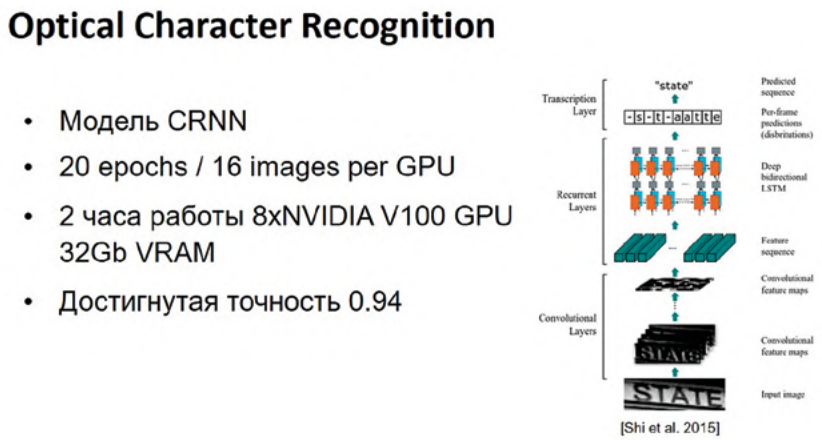


Рис. 3. Характеристики модели распознавания символов тибетской письменности.

В настоящее время требуется разработка полнофункциональных моделей распознавания символов тибетской и монгольской письменностей, машинного перевода с тибетского и старописьменного монгольского языков. Срок реализации проекта — 4 года. Полное финансирование проекта — 70 млн руб.

Преимущества и уникальность

Уникальная источниковая база исследований и междисциплинарный подход с широким применением технологий искусственного интеллекта обеспечат высокую теоретическую значимость и практическую ценность результатов, соответствующих мировым научным достижениям в этой области. Выполнение проекта позволит утвердить приоритет российской науки в данной области, сформировать передовой теоретический и практический задел для внедрения методов искусственного интеллекта и других цифровых технологий в сфере сохранения, изучения и освоения историко-культурного наследия, дать новый импульс развития российскому востоковедению.

Контекст

Буддийское наследие, особенно канонические собрания, исторически играет важную структурообразующую роль для социально-политической жизни в традиционных и современных азиатских обществах. Оно выступает ключевым фактором непрерывности передачи знаний и духовных традиций, сохранения культурной и национальной идентичности. Ввиду этой особой социальной роли буддийские книжные собрания выступали на протяжении истории важнейшими объектами приложения интеллектуальных усилий и технологических новшеств. Обращение к буддийскому наследию, его актуализация всегда происходили в период глубоких социальных трансформаций. Эта традиция сохраняется до настоящего времени. Так, модернизация начала XX века в Японии ознаменовалась появлением инновационного Канона Тайсё (Тайсё Дайдзокё). Новые редакции китайского буддийского канона предпринимались в Китайской республике в 1930-е, в КНР и на Тайване в 1970–80-е годы. С начала 1990-х годов правительство КНР предприняло ряд масштабных проектов по сохранению и изучению культурного наследия национальных меньшинств, в рамках одного из которых в 2002–2010 годах было выполнено 400-томное факсимильное издание монгольских буддийских канонических собраний Ганджур и Данджур.

С развитием цифровых технологий в последние десятилетия активизировалась деятельность по сохранению и изучению буддийского письменного наследия, представленного на разных языках. В этой области успешно функционируют международные консорциумы, такие как Asian Classics Input Project (ACIP), Buddhist Digital Resource Center (BDRC), Chinese Buddhist Electronic Text Association (CBETA) и др., которые предлагают обширные цифровые ресурсы, раскрывающие многоязычное письменное наследие буддийской цивилизации. Наиболее значительных успехов в цифровизации письменного наследия тибетского буддизма достиг консорциум BDRC, партнером которого выступает Служба цифрового хранения (DRS) Гар-

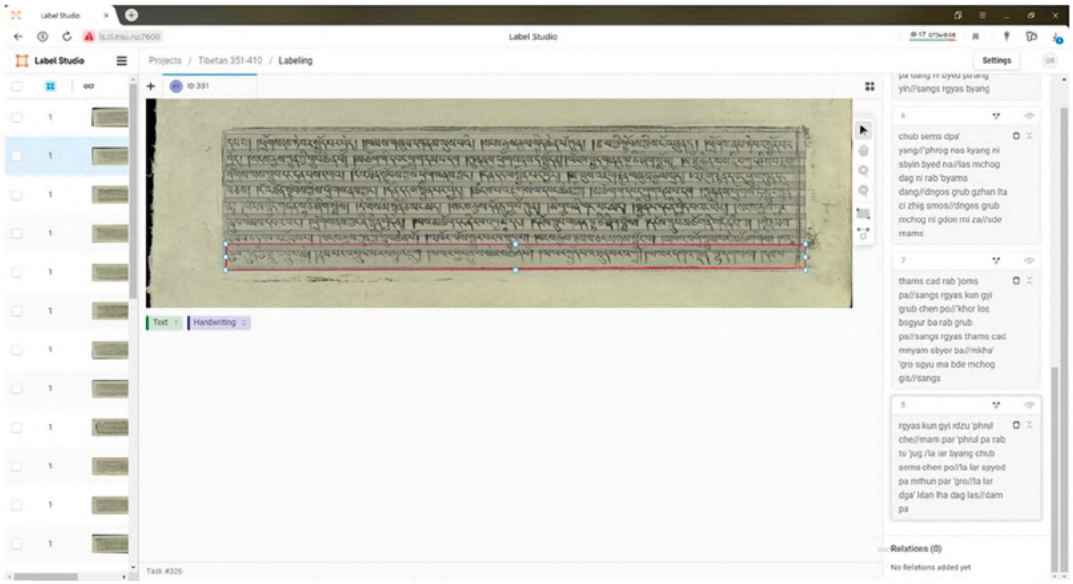


Рис. 4. Пример датасета тибетского текста в приложении LabelStudio.

вардского университета. К настоящему времени создаваемая совместными усилиями цифровая библиотека тибетских текстов стала важным ресурсом для тибетологии и буддологии. Каждый год к ресурсам BDRC обращаются из 150 стран мира [Access 2012]. BDRC оцифровал значительное число буддийских сочинений, включая редкие рукописные и ксилографические издания, полученные от партнерских организаций, которые включают Библиотеку Конгресса США, Библиотеку тибетских трудов и архивов (LTWA) в Дхарамсале (Индия), управляемую Центральной тибетской администрацией, монастырские и университетские библиотеки из разных стран, где хранятся памятники тибетской литературы. Помимо сканированных книг, представленных аннотированными цифровыми изображениями, BDRC занимается обработкой полнотекстовых материалов, представленных в машиночитаемых форматах, которые поступают из нескольких источников, в частности созданных в рамках программы Namsel OCR (оптического распознавания символов), реализуемой в Калифорнийском университете в Беркли. Размещаемые материалы индексируются по нескольким категориям, включая обширные предметные таксономии, персоны, места. Благодаря этому создается глубокое представление сложного и многогранного мира тибетской литературы.

В 2021 году BDRC запустил новую цифровую платформу Buddhist Digital Archives (BUDA), которая предоставляет доступ к обширной коллекции произведений тибетского буддизма и разнообразным материалам на санскрите, китайском, пали, бирманском и кхмерском языках, с которыми BDRC работает с 2016 года. Новая платформа построена на принципах открытого исходного кода. В ее основе лежат две основные архитектуры, International Image Interoperability Framework (IIIF) и Linked Open Data (связанные открытые данные, LOD). Стандарты IIIF применяются для управления цифровыми изображениями и их отображения. Архитектура LOD предназначена для структурирования машиночитаемых данных, что облегчает объединение, извлечение и обновление сложноорганизованных массивов информации на семантическом уровне. Через механизмы LOD налажен обмен данными между BUDA и базой данных проекта Sakya Research Centre (SRC) Австрийской академии наук, в которой собран большой объем авторитетных данных (персоналий, географических локаций и т.д.), связанных с тибетской культурой.

Обращение цифровых гуманитарных проектов и библиотек к технологиям, подобным IIIF и LOD, обусловлено тем, что их внедрение позволяет интегрировать ранее разрозненные цифровые хранилища материалов, связанных с культурным наследием, в обширную сеть семантически взаимосвязанных ресурсов и формировать на этой основе глобальное информационное пространство. Несмотря на значительные усилия, предпринимаемые в основном в рамках инициатив ACIP и BDRC, объем полнотекстовых материалов,

Ринчинов Олег Сергеевич

Личный комментарий от разработчика:

«Буддийское письменное наследие, канонические книжные собрания играют структурообразующую роль в социально-политической жизни азиатских стран. Цифровая библиотека послужит важным мостом сотрудничества и культурного обмена между Россией и странами буддийской цивилизации, наиболее динамично развивающимися в сегодняшнем мире».

связанных с наследием тибето-монгольского буддизма, и переводов письменных памятников на европейские языки продолжает оставаться незначительными. Острыми остаются проблемы критического исследования, перевода и интерпретации первоисточников буддийской классики, например разновременных изданий канонических собраний Кангьюр и Тенгьюр. Для того чтобы представить буддийское наследие Центральной Азии и Байкальского региона во всей его полноте, возникает настоятельная необходимость во внедрении новых методов в цифровизации буддийского письменного наследия, его изучении и освоении, разработке эффективных методов анализа и систематизации информации, вводе в научный оборот новых источников.

Дальнейшее развитие в этой области связано с развитием методов оптического распознавания символов (OCR) восточных письменностей и машинного перевода. Однако существующие на сегодня решения, например упомянутый выше Namsel OCR, способные достаточно результативно обрабатывать печатные тексты (книги современного формата), остаются малопригодными для рукописей и ксилографов, которые отличаются неоднородным качеством воспроизведения текста, плохой сохранностью бумаги и т.д. Для монгольской классической письменности цифровые решения в этой области отсутствуют в принципе. Поэтому для успешного и эффективного решения указанных задач представляется актуальным внедрение технологий искусственного интеллекта, таких как нейросетевые алгоритмы «глубокого обучения», для распознавания символов тибетской и монгольской письменности и машинного перевода, развитие корпусных методов текстологического анализа на основе подходов big data и других современных цифровых инструментов.

Предполагаемый интерес для внедрения

Реализация проекта обеспечит переход к передовым технологиям сохранения и изучения важной части культурного наследия, его актуализации в современных академических исследованиях и общественной жизни. Его ценность заключается в многосторонней реконструкции буддийского культурного пространства в контексте исторически сложившихся тесных связей восточных регионов России и буддийской Азии и формировании новой источниковой базы для переосмысления существующих и получения новых научных знаний на основе цифровизации уникальных памятников письменного наследия. С использованием методов машинного обучения создаются и внедряются средства оптического распознавания символов тибетской и монгольской письменностей и системы машинного перевода с тибетского и старописьменного монгольского языков, что позволит сформировать и ввести в публичный оборот обширный цифровой корпус произведений буддийской классики. В его составе будут представлены уникальные издания буддийского канона на тибетском и монгольском языках, ранее неизвестные и малоизученные памятники письменного наследия Центральной Азии и Байкальского региона.

Выполнение проекта позволит консолидировать научные кадры в ИМБТ СО РАН, интенсифицировать сотрудничество с научными учреждениями России и зарубежной Азии. Результаты проекта будут востребованы российскими и зарубежными буддийскими общинами, научными учреждениями, они будут способствовать позиционированию России в культурном и религиозном пространстве стратегически важного региона Восточной и Центральной Азии.

НАБОР ДАННЫХ EMNIST TCYRILLIC ДЛЯ ТРЕНИРОВКИ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ

РАСПОЗНАВАНИЯ СИМВОЛОВ ТЮРКСКИХ ПИСЬМЕННОСТЕЙ, ПОСТРОЕННЫХ НА ОСНОВЕ РАСШИРЕННОЙ КИРИЛЛИЦЫ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Хакасский государственный университет им. Н.Ф. Катанова», Инженерно-технологический институт
655017, Республика Хакасия, г. Абакан, проспект Ленина, 92/1

Голубничий Артем Александрович
старший преподаватель кафедры программного обеспечения вычислительной техники и автоматизированных систем

Яблонцева Арина Дмитриевна
студент кафедры программного обеспечения вычислительной техники и автоматизированных систем

Голубничий Артем Александрович, старший преподаватель кафедры программного обеспечения вычислительной техники и автоматизированных систем

Яблонцева Арина Дмитриевна, студент кафедры программного обеспечения вычислительной техники и автоматизированных систем

В области искусственного интеллекта и машинного обучения существует множество наборов данных, используемых для тренировки моделей машинного обучения, наиболее популярными из них являются MNIST (набор рукописных символов), CIFAR-10 (набор тематических изображений из 10 классов), ImageNet (набор аннотированных изображений с несколькими фактами об объекте) и ряд других датасетов, относящихся к конкретной предметной области — такие как Fashion-MNIST (изображения одежды), Caltech-UCSD Birds 200 (изображения птиц) и т.д. При этом в настоящее время отсутствуют достаточно серьезные и массивные датасеты, содержащие сведения о печатных буквах основной и расширенной кириллиц, содержащих множественные искажения.

Актуальность создания такого датасета для детальной настройки моделей машинного обучения обусловлена тем, что в настоящее время культурное наследие многих народов России, в том

а б в г д е ё ж з и й к л м н о п р с т у
х ц ч ш щ ъ ы ь э ю я ө ӊ ғ ү ә ҥ ӄ ӧ ӱ і
ҷ ј ў ҫ ҳ ӈ ҙ ӄ ҡ ҥ ҰҰ ҵ ҹ ҝ ӑ ӓ ӗ ӳ А Б В
Д Е Ё Ж З И Й К Л М Н О П Р С Т У Ф Х Ц Ч
Щ Ъ Ы Ь Э Ю Я Ө Ӊ Ғ Ү Ә Ҥ Ӄ Ӧ Ӱ І Ҹ Ҷ Ј Ў
Ҳ Ӈ Ҙ Ӄ Ҡ Ҥ ҰҰ Ҵ Ҹ Ҝ Ӑ Ӓ Ӗ Ӳ ’ 1 2 3 4 5 6
8 9 0 ! ? . , : ; " № ( ) - '

Рис. 1. Дизайнер процессов.

числе Енисейской Сибири, представлено в виде печатных трудов, вышедших в период 1930–1980 годов. Твердые копии многих произведений находятся в достаточно ветхом состоянии, использование стандартных средств оцифровки без осуществления распознавания символов в значительной степени усложняет использование исходных текстов. При этом языковые модели для систем распознавания символов разработаны исключительно для популярных языков. Таким образом, создание набора датасетов для обучения моделей машинного обучения в разном количестве классов имеет значительную актуальность.

Исходный набор данных EMNIST TCyrillic составляет 74 088 изображений, относящихся к одному из 147 классов. Таким образом, количество изображений на один класс составляет всего 504 экземпляра. Столь малый объем вполне приемлем для печатных наборов данных по причине того, что исполнение букв в печатном виде несет значительно меньший уровень вариаций и искажений в сравнении с рукописным начертанием. Соответственно отпадает необходимость в использовании нескольких тысяч изображений для каждого класса, как это представлено в наборе MNIST (более 7000 изображений каждого символа).

Основная цель создания данного датасета сводится к созданию набора данных, позволяющих получить схожие значения предсказания в сравнении с классическими наборами данных и методами предсказания.

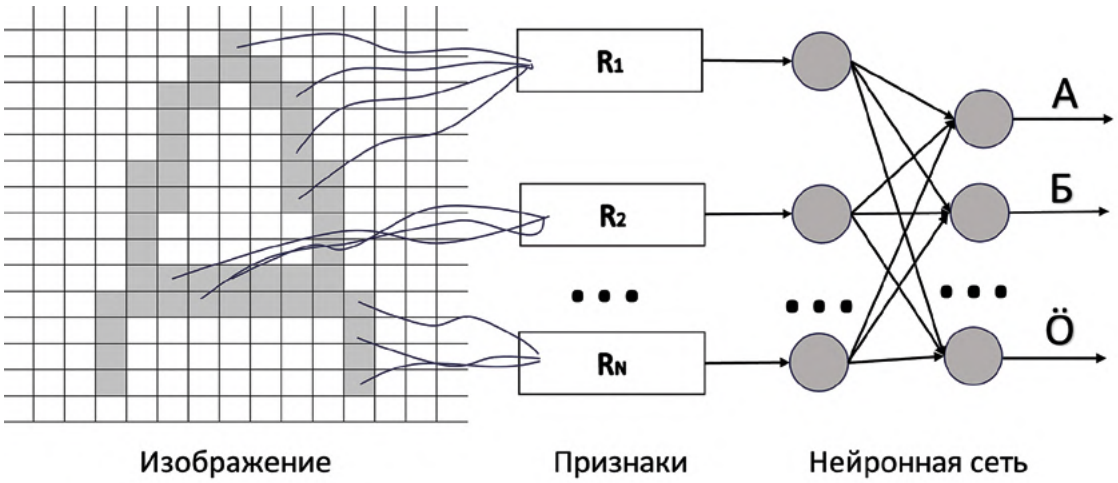


Рис. 2. Логика извлечения признаков для классификации символов.

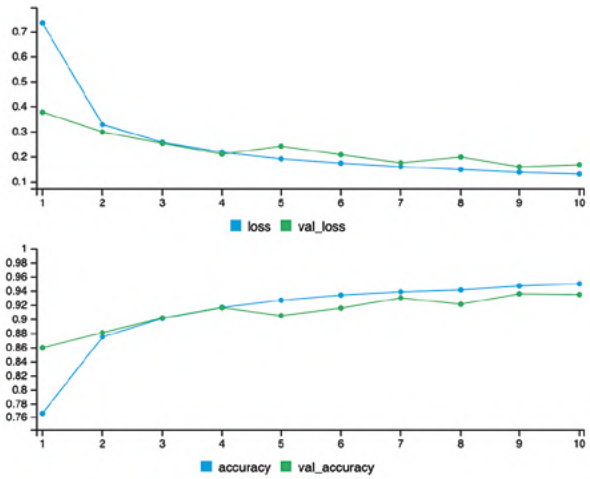


Рис. 3. Результаты предсказания для набора данных MNIST.

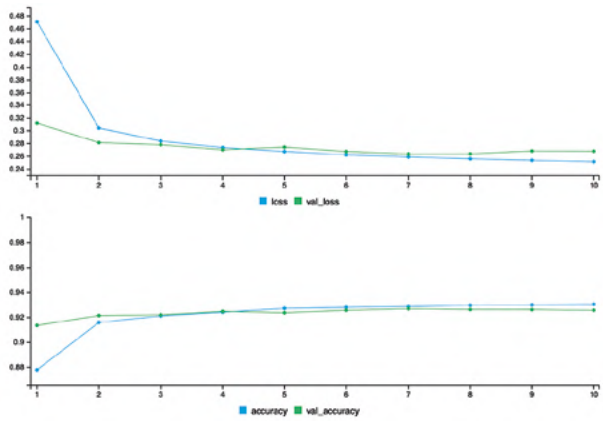


Рис. 4. Результаты предсказания для набора данных EMNIST TCyrillic

Первые результаты для набора данных при использовании линейного классификатора, реализованного посредством персептрона, дает соразмерный уровень предсказания на уровне 92–93%, что соответствует аналогичным результатам для набора данных MNIST (92%).

На текущем этапе разработан набор данных, проводится измерение точности результатов и сравнение с базовым набором данных MNIST. В качестве дополнительных доработок планируется изменение размерности набора данных в зависимости от точности распознавания символов.

Набор данных является уникальным с точки зрения решаемых задач. Так, в настоящее время отсутствуют датасеты, цель которых — тренировка моделей машинного обучения для оптического распознавания символов расширенной кириллицы.

Данный продукт может быть внедрен для построения систем оптического распознавания символов с целью оцифровки архивов, содержащих письменность на расширенной кириллице тюркских народов.

Социальные эффекты: наборы данных можно использовать для ускорения оцифровки национальных текстов, в том числе малых народов, что будет способствовать сохранению и развитию культуры народов России.

Выступление на форуме «Наука будущего – наука молодых» г. Орел 2023.

НЕЙРОСЕТЕВАЯ СИСТЕМА ДИАГНОСТИКИ ОСТРОГО ИНСУЛЬТА

НА ОСНОВЕ КОМПЬЮТЕРНОГО АНАЛИЗА КТ-ИЗОБРАЖЕНИЙ ГОЛОВНОГО МОЗГА

Федеральное государственное бюджетное учреждение науки Институт математики им. С. Л. Соболева Сибирского отделения Российской академии наук
630090, г. Новосибирск, пр. Академика Коптюга, 4

Бериков Владимир Борисович
доктор технических наук, доцент, главный научный сотрудник
тел.: (383) 329-75-75, berikov@math.nsc.ru

Краткое описание разработки

В институте математики имени С. Л. Соболева СО РАН разработана нейросетевая система диагностики острого инсульта с помощью анализа изображений бесконтрастной компьютерной томографии головного мозга. В создании системы участвовали сотрудники Международного томографического центра СО РАН, ФИЦ ИВТ, аспиранты и студенты Новосибирского государственного университета.

Разработанная система искусственного интеллекта предназначена для автоматизированной постановки диагноза (наличие поражения, его тип, локализация очага) и может служить «цифровым помощником» врача-рентгенолога. Важно отметить, что бесконтрастное сканирование, хотя и дает меньше информации, чем другие варианты томографического исследования, имеет значительно меньше противопоказаний и менее затратно. Эти особенности особенно значимы для оперативной диагностики и назначения лечения.

В основе архитектуры системы лежит сверточная сеть глубокого обучения U-Net. В ходе работы были также созданы и экспериментально исследованы различные варианты нейросетевой модели с применением оригинальных модификаций базовой архитектуры.

Для обучения нейронной сети был использован созданный в МТЦ СО РАН анонимизированный банк трехмерных бесконтрастных КТ-изображений головного мозга у 100 больных при остром инсульте. Изображения были вручную размечены рентгенологами: на каждом срезе, содержащем пораженные участки, были очерчены контуры очагов поражения. Для устранения различного рода неточностей разметки привлекались несколько специалистов, параллельно обрабатывающих снимки.

На тестовой выборке проведено оценивание качества нейросети. Точность решений, принятых нейронной сетью, оказалась сопоставима с результатами, полученными специалистами-рентгенологами при значительно более высокой скорости обработки изображений (несколько секунд против нескольких минут). Пример получаемых решений показан на рис. 1. Точность локализации патологического очага ишемического инсульта моделью составила около 70% по метрике Dice Similarity Coefficient (DSC); сбалансированная точность (среднее показателей чувствительности и специфичности) распознавания наличия поражения на снимке — 87,5%.

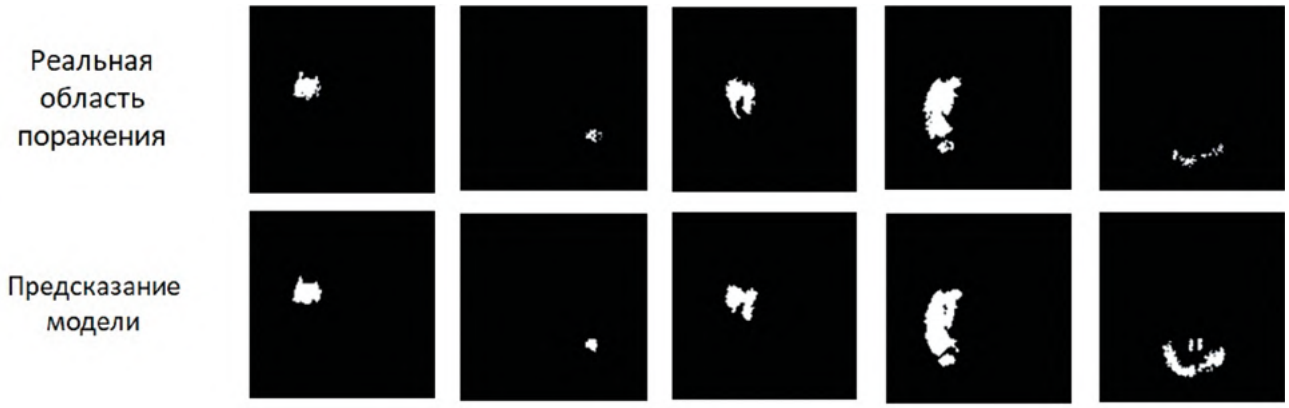


Рис. 1. Пример результатов работы системы: верхний ряд – маски реальных областей поражения, нижний ряд – маски, полученные нейросетевой моделью.

Исследование показало: значительная доля ошибок нейросетевой модели приходится на зоны поражения с небольшим объемом. При этом обнаружено, что если ставить задачей прогноз областей поражения с объемом выше определенного порога, то точность сегментации существенно улучшается. Пациенты с более обширной зоной поражения нуждаются в первоочередном лечении, поэтому их выявление является важной задачей.

Сеть помогает принимать медицинские решения при большом потоке пациентов, отсеивая «нормальные» случаи от пациентов с пораженными участками. Таким образом, специалисты получают возможность в первую очередь осматривать больных, нуждающихся в неотложной помощи.

Преимущества и уникальность

Известные методы построения глубоких нейросетей имеют ряд ограничений, таких как значительная ресурсоемкость алгоритмов, трудность интерпретации решений. Предлагаемые решения ориентированы на преодоление указанных трудностей. Для повышения степени интерпретируемости нейросетевых решений разработана архитектура сети, в которой используется оригинальный подход, основанный на комбинации сверточной нейронной сети и нечеткого дерева решений по сходству. Метод позволяет получать интерпретируемые результаты анализа на языке логических суждений о сходстве областей изображения с некоторыми эталонами.

В настоящее время активно развиваются методы и алгоритмы анализа медицинских изображений, в частности, снимков компьютерной томографии. Для этого применяются глубокие сверточные нейросети, а также «классические» подходы компьютерного зрения. Существует несколько отечественных и зарубежных аналогов разработанной системы, основанных на сверточных нейронных сетях, обладающих схожим функционалом. Часть разработок предназначена для решения задачи распознавания наличия инсультного поражения на КТ снимке, другая часть позволяет дополнительно определить место локализации очага. В ряде стран существуют специализированные бригады скорой помощи (Mobile Stroke Units), где могут использоваться автоматизированные системы бесконтрастной КТ диагностики острого инсульта.

Разработка может быть интересна для разработчиков программного обеспечения, предназначенного для анализа изображений компьютерной томографии. Приветствуются различные формы индустриального партнерства. Система позволит автоматизировать и ускорить постановку диагноза острого инсульта по бесконтактным КТ-изображениям и может служить «цифровым помощником» врача-рентгенолога.

Аналог разработанной системы, основанный на сверточных нейронных сетях

ИНТЕЛЛЕКТУАЛЬНАЯ ТЕХНОЛОГИЯ ПРОГНОЗИРОВАНИЯ ИНСОЛЯЦИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Хакасский государственный университет им. Н.Ф. Катанова», Инженерно-технологический институт
655017, Республика Хакасия, г. Абакан, проспект Ленина, 92/1

Логотип интеллектуальной технологии прогнозирования инсоляции на основе оригинальной модифицированной нечеткой нейросети (МНН)

Энгель Екатерина Александровна
кандидат технических наук, доцент, доцент кафедры цифровых технологий и дизайна Инженерно-технологического института ФГБОУ ВО «ХГУ им. Н.Ф. Катанова»
тел.: +7-913-051-64-52, Angel@khsu.ru, ekaterina.en@gmail.com

Разработка автоматически формирует макет «Рынок на сутки вперед» для солнечных электростанций.

Интеллектуальная технология прогнозирования инсоляции на основе оригинальной модифицированной нечеткой нейросети (TRL2) предназначена для солнечных электростанций, обеспечивает снижение суммы штрафов для участников рынка на сутки вперед при формировании почасового макета выработки электроэнергии.

Разработанная интеллектуальная технология прогнозирования инсоляции, интегрированная с разрабатываемой интеллектуальной технологией прогнозирования выработки электроэнергии солнечной электростанцией обеспечит формирование эффективного почасового макета выработки электроэнергии при погрешности прогноза погоды в среднем не более 1%.

Преимущества разработки перед аналогами

Интеллектуальная технология непрямого прогнозирования инсоляции на основе оригинальной модифицированной нечеткой нейросети настраивается на архивных данных, автоматическая самоадаптация и самообучение (нейросетевые технологии); интеллектуальная технология прогнозирования выработки электроэнергии солнечной электростанции автоматически формирует макет «Рынок на сутки вперед».

Отрасли и области применения результатов исследования соответствуют Стратегии социально-экономического развития Республики Хакасия, поскольку регион обладает высоким природным потенциалом для развития солнечной энергетики (величина годовой сол-

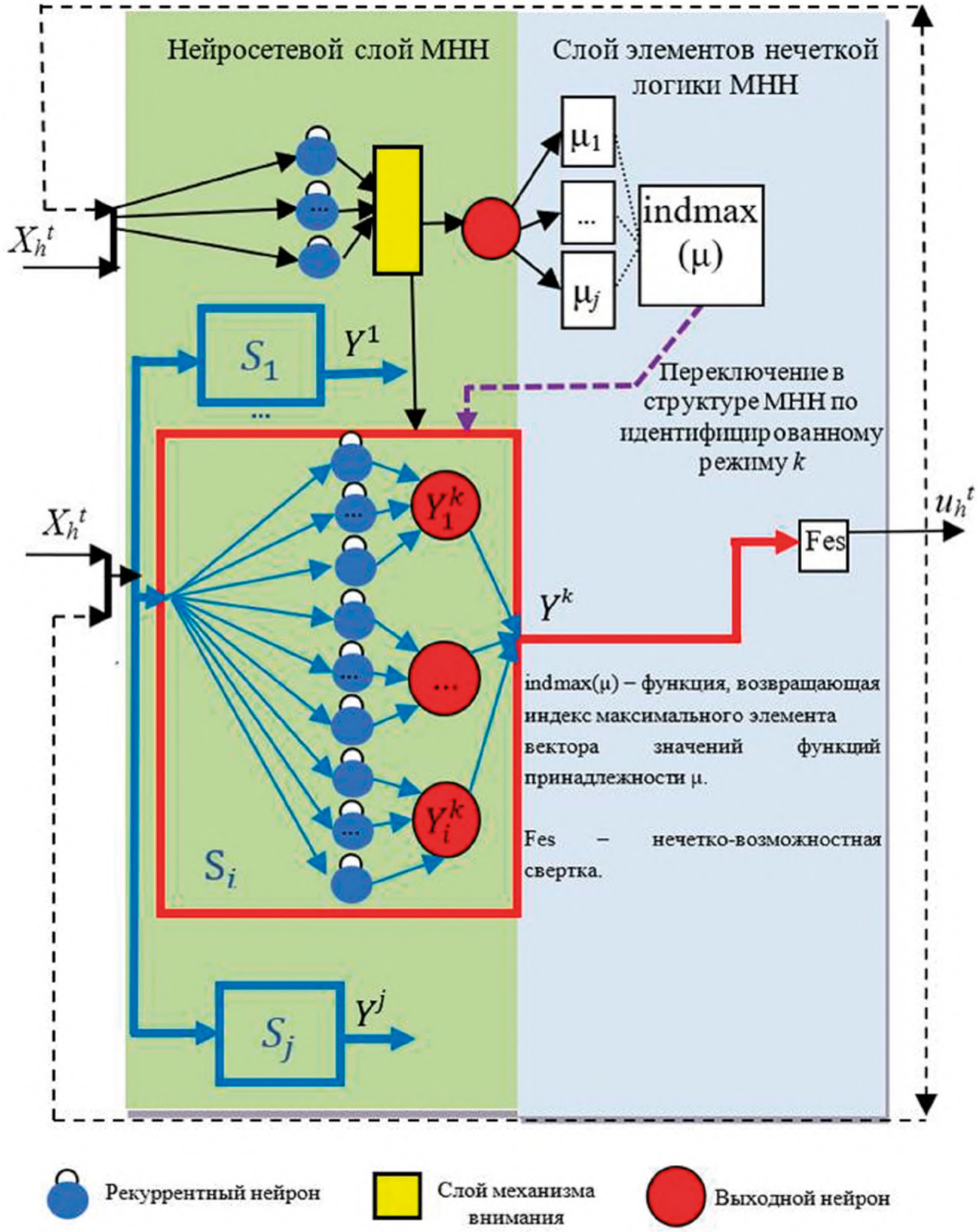
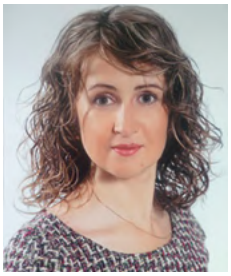


Рис. 1. Схема интеллектуальной технологии прогнозирования инсоляции на основе оригинальной модифицированной нечеткой нейросети (МНН)



Энгель Екатерина Александровна

Личный комментарий от разработчика:

«Разработанная интеллектуальная технология прогнозирования инсоляции, интегрированная с разрабатываемой интеллектуальной технологией прогнозирования выработки электроэнергии солнечной электростанцией, достаточно проста, не требует перепрошивки инверторов солнечной электростанции и больших затрат, легко интегрируется в информационную систему солнечной электростанции».

нечной инсоляции соответствует, например, уровню инсоляции в южных районах Краснодарского края). Результаты исследования развивают интеллектуальные технологии возобновляемой электроэнергетики в Сибири и Российской Федерации, способствуют снижению выбросов парниковых газов, улучшению экологической обстановки и экологического благополучия населения. Преимущества разработки перед существующими современными аналогами, в том числе нейросетевыми: интеллектуальная технология непрямого прогнозирования инсоляции на основе оригинальной модифицированной нечеткой нейросети автоматически настраивается на архивных данных, присутствуют автоматическая самоадаптация и самообучение.

Предполагаемый интерес для внедрения: энергетика, солнечные электростанции.

Запрос на индустриальное партнерство

Внедрение технологии на солнечной электростанции, покупка лицензии. Мероприятия по развитию цифровых сервисов для НИОКТР, работе с Big Data и суперкомпьютерными вычислениями в коллаборации с солнечными электростанциями. Создание цифрового сервиса для упрощения работы с данными, предоставление доступа клиентам, применение оригинальной модифицированной нечеткой нейросети при анализе и представлении данных.

Финансирование

Исследование выполнено за счет средств гранта министерства образования и науки Республики Хакасия (Соглашение № 91 от 13.12.22).

ТЕХНОЛОГИЯ ДЛЯ ПЕРСОНАЛИЗИРОВАННОЙ ДИАГНОСТИКИ НЕЙРООНКОЛОГИЧЕСКИХ ЗАБОЛЕВАНИЙ

С ПОМОЩЬЮ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Федеральное государственное автономное образовательное учреждение высшего образования «Новосибирский национальный исследовательский государственный университет»
630090, г. Новосибирск, ул. Пирогова, 2

Летягин Андрей Юрьевич
доктор медицинских наук, профессор, заместитель директора НИИКЭЛ — филиала ИЦиГ СО РАН

Амелин Михаил Евгеньевич
кандидат медицинских наук, руководитель отделения ФГБУ «ФЦН» Минздрава России (г. Новосибирск)

Амелина Евгения Валерьевна
кандидат физико-математических наук, научный сотрудник ЛАПДиМО ММФ НГУ

Тучинов Баир Николаевич
младший научный сотрудник ЛАПДиМО ММФ НГУ

Павловский Евгений Николаевич
кандидат физико-математических наук, ведущий научный сотрудник ЛАПДиМО ММФ НГУ

Голушко Сергей Кузьмич
доктор физико-математических наук, доцент, профессор кафедры ММод ММФ НГУ

Использование алгоритмов и методов ИИ в качестве помощника врачу и пациенту

В соответствии с указом Президента Российской Федерации «О Стратегии научно-технологического развития Российской Федерации» от 1 декабря 2016 г. N 642 одним из основных приоритетов государственной политики в сфере здравоохранения является «...переход к персонализированной медицине, высокотехнологичному здравоохранению и технологиям здоровьесбережения, в том числе за счет рационального применения лекарственных препаратов (прежде всего антибактериальных)».

Данная задача может быть решена с помощью современных технологий искусственного интеллекта: мобильной телемедицины, экспертных систем, медицинских чат-ботов, систем по анализу генома, инструментов для анализа медицинских изображений, машинного обучения и нейронных сетей, систем поддержки принятия врачебных решений и т.д.

Высокий уровень смертности от онкологических заболеваний в РФ (более 200 человек на 100 тысяч населения) прежде всего связан с поздней диагностикой заболевания. Поэтому вызовами для современной науки являются: раннее выявление, дооперационная диагностика гистологического типа опухоли и ее характеристик, персонализированное лечение и оценка рисков заболеваний и формирование персона-

Тучинов Баир Николаевич

Павловский Евгений Николаевич

Летягин Андрей Юрьевич

Голушко Сергей Кузьмич

Амелин Михаил Евгеньевич

Амелина Евгения Валерьевна

лизированных рекомендаций по управлению этими рисками. Работы в этих направлениях ведутся как коммерческими организациями (Google, IBM, Nvidia), так и мировыми научно-исследовательскими и образовательными учреждениями, что говорит о крайней важности и сложности перечисленных задач.

Основной причиной для разработки и внедрения технологий искусственного интеллекта (ИИ) в нейроонкологии является высокая распространенность опухолей головного мозга — до 200 случаев на 100 тысяч населения. Частота встречаемости первичного очага в головном мозге составляет 5–10%, но у 60–70% умерших от злокачественных новообразований обнаруживаются метастазы в головном мозге.

Магнитно-резонансная томография — наиболее распространенный метод первичной неинвазивной диагностики опухолей головного мозга и контроля динамики заболевания. Одними из самых сложных задач в этой области являются классификация типов опухолей и определение клинических параметров (размер и объем) для проведения диагностики и лечебных процедур, в том числе операции.

Развитие МРТ-технологий революционно изменило диагностику, лечение, контроль и прогнозирование заболеваний головного мозга. Однако «ручная» обработка изображений требует много времени и не исключает ошибок в интерпретации: уровень ежедневных ошибок и расхождений в радиологии превышает 3–5%.

Предлагаемая разработка нацелена на решение проблемы визуализационного анализа для персонализированной диагностики и выработки рекомендаций по лечению нейроонкологических заболеваний (опухолей головного мозга) с применением методов нейровизуализации.

Разработчиком технологии является научный коллектив высококвалифицированных специалистов различных направлений: лучевых диагностов, неврологов, нейроонкологов, математиков, специалистов ИИ и программистов.

Размеченный специалистами набор данных Siberian brain tumor dataset (SBT), включающий 1000 серий снимков, в настоящий момент содержит дооперационные снимки для четырех типов первичных опухолей: астроцитомы (65), глиобластомы (105), менингиомы (150), невриномы (125) и вторичных — метастазов (85). Остальная часть базы — послеоперационные МР-томограммы к соответствующим дооперационным снимкам. Все данные в базе анонимизированы.

На основе систем искусственного интеллекта (глубокого машинного обучения) разработаны алгоритмы и модели:

- мультиклассовой сегментации опухоли головного мозга с помощью предварительной обработки данных комбинаций МРТ-последовательностей;

РЕЗУЛЬТАТЫ СЕГМЕНТАЦИИ НА SBT

МОДЕЛИ	DiceET	DiceTC	DiceWT	Sens	Spec
LinkNet	0,826	0,826	0,911	0,872	1
nnU-net	0,846	0,867	0,917	0,881	1
DMFNet	0,796	0,827	0,9	0,851	1
TransBTS	0,756	0,753	0,858	0,868	1

- сегментации опухолей головного мозга и оценки связанной с ней неопределенности на глиомах высокой и низкой степени злокачественности;
- классификации опухолей головного мозга с использованием методов искусственного интеллекта, основанных на внимании (attention).

Реализован протокол автоматической сегментации опухолей головного мозга в виде приложения для кроссплатформенной системы обработки и визуализации медицинских изображений с открытым исходным кодом 3D Slicer®. В основе программного модуля лежат двух- и трехмерные модели компьютерного зрения с предварительной обработкой данных МРТ-последовательностей: предконтрастное T1-взвешенное изображение, постконтрастное T1-взвешенное изображение, T2-взвешенное изображение, T2-взвешенное изображение с технологией инверсии-восстановления с ослаблением сигнала от жидкости.

Оно позволяет обнаруживать и распознавать 4 типа опухолей головного мозга (менингиома, невринома, глиобластома, астроцитома), а также сегментировать и выделять компоненты опухоли (некроз, поглощающая Gd-контраст опухолевая ткань, непоглощающая Gd-контраст опухолевая ткань, перитуморальный отек) и оценивать их размеры.

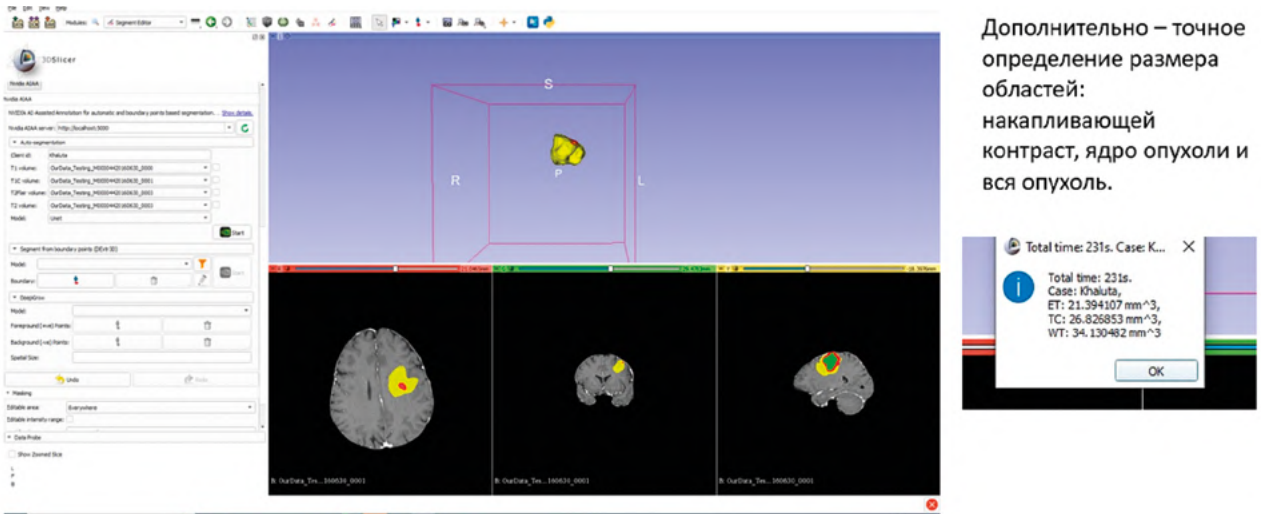


Рис. 1. Пример автоматической сегментации.

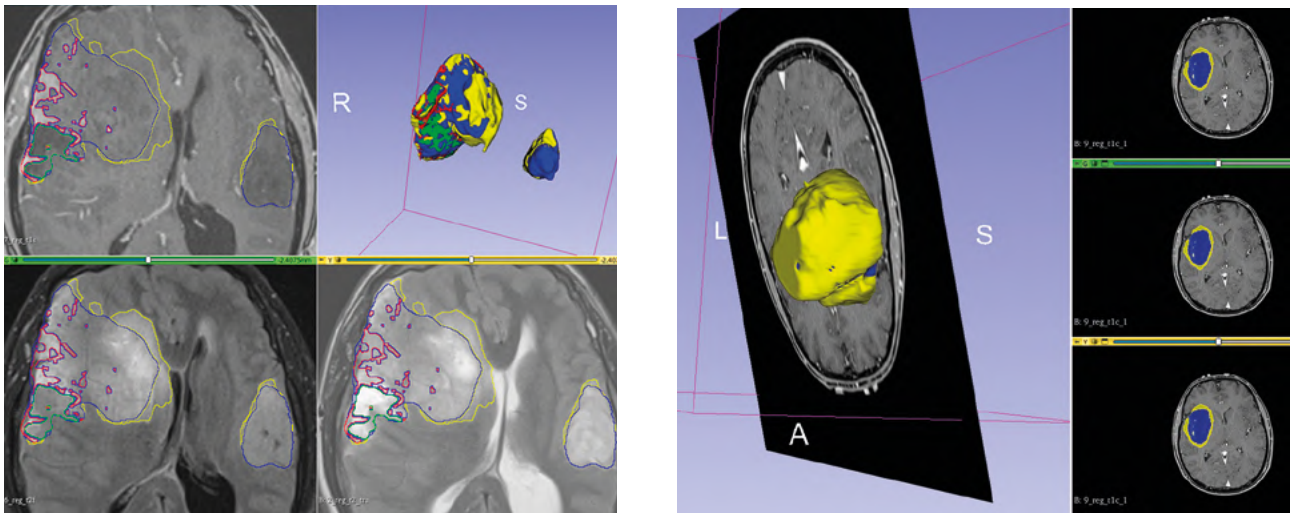


Рис. 2. Пример визуалиализации глиобластомы. Аннотации включают метки перитуморального отека (желтый), неконтрастируемой опухоли (синий), GD-контрастируемой опухоли (красный) и ядра некротической опухоли (зеленый).

Уровень готовности технологии — TRL- 5.

Разработанная и апробированная технология (системы) для диагностики опухолей головного мозга с помощью искусственного интеллекта имеет следующие преимущества и возможности:

1. Производительность. Ежегодно проводится в РФ более 1 млн исследований МРТ головного мозга. Расшифровка каждого занимает 2–3 часа, в более серьезных случаях до 24 часов. Время обработки системы — 10–15 минут.

2. Возможности использования в качестве:

   - аналитического инструмента специалиста-радиолога для идентификации, сегментации и классификации опухолей ЦНС;

   - обучающей выборки для разработки экспертных систем на основе ИИ;

   - базы данных для проведения исследовательских работ в области радиономики;

   - обучающего курса для специалистов-радиологов, специалистов по искусственному интеллекту, а также для разработки и выполнения проектов по распознаванию заболеваний, диагностируемых с помощью МРТ-визуализации.

Полученные результаты — объекты интеллектуальной собственности.

Свидетельство о регистрации ноу-хау № 48 от 20.08.2021 «Система для автоматической сегментации опухолей головного мозга» Новосибирского государственного университета.

Программа «Система для дифференциальной диагностики опухолей головного мозга на МРТ-изображениях». Получено Свидетельство о государственной регистрации программы для ЭВМ № 2022685266. Правообладатель: федеральное государственное автономное образовательное учреждение высшего образования «Новосибирский национальный исследовательский государственный университет» (RU).

Возможность применения — медицинские учреждения, клиники и медицинские центры.

Полученные результаты могут использоваться для обучения медицинских специалистов применению ИИ в здравоохранении.

Находимся в поиске партнёра для проведения дополнительных НИОКР с целью разработки прототипов продукции (технологии) для демонстрации потенциальным инвесторам, внедрения и повышения TRL-разработки и вывода на рынок.

МЕТОД ОЦЕНКИ ЭФФЕКТИВНОСТИ ПРИМЕНЕНИЯ ГЕОФИЗИЧЕСКИХ ИССЛЕДОВАНИЙ

ДЛЯ РЕШЕНИЯ ГЕОЛОГОРАЗВЕДОЧНЫХ ЗАДАЧ, ОСНОВАННЫЙ НА ПРИМЕНЕНИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И ПОСТРОЕНИИ ТРЕХМЕРНОЙ ГЕОЛОГО-ГЕОФИЗИЧЕСКОЙ МОДЕЛИ

ООО «Стрекоза» Академпарк
новосибирского Академгородка
630090, г. Новосибирск,
ул. Николаева 12, 77

Иванова Ирина Игоревна
геофизик-эксперт
admin@fly-dragon.ru

О прогнозных геолого-геофизических моделях

Прогнозные геологические модели на основе интерпретации результатов геофизических исследований обычно строятся методом комплексирования данных измерений, что основано на зависимости различных физических свойств, то есть результатов измерений, от типа горных пород и руд.

Процесс построения геологической модели творческий (авторский, субъективный), поэтому нередко при работе различных специалистов получаются различные модели, что обосновывает необходимость оценки качества прогноза и выбора наиболее подходящего решения.

Мы подошли к процессу выбора оптимальной модели с точки зрения оценки риска недостоверности прогноза и выбора наиболее подходящих методов снижения неопределенности в составе и структуре построенной геолого- геофизической модели.

Предлагается использовать цифровой геолого-геофизический двойник участка работ и метод оценки эффективности геофизических работ на основе машинного обучения. Искусственный интеллект активно используется для интерпретации данных сейсмических и петрофизических исследований в нефтяной геофизике [1], а также в работах с ГИС [2]. Для исследований зависимостей между физическими параметрами, полученными в ходе полевых экспериментов, и геологическим типом пород обычно используют статистические методы.

Мы предлагаем решение, основанное на ИИ, позволяющее делать выводы об эффективности

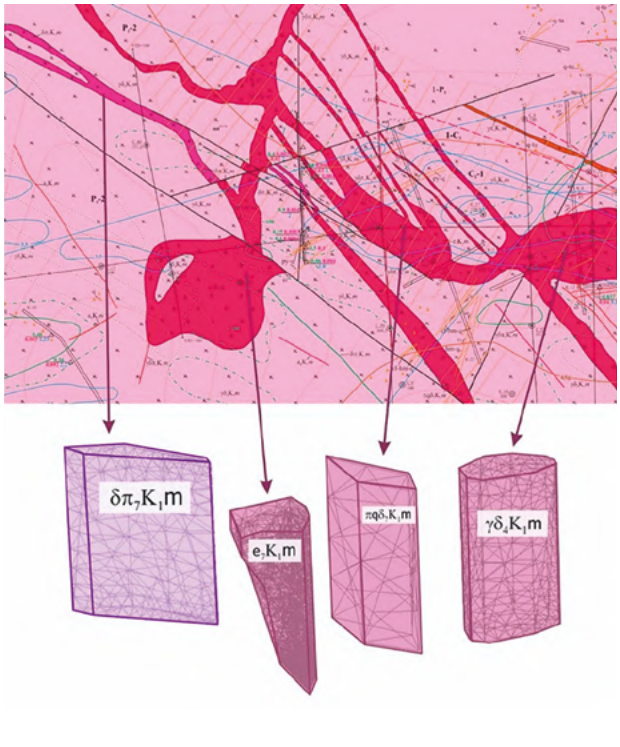


Рис. 1. Преобразование картографических объектов в сеточные, внесение соответствующих им физических свойств в каталог.

применения разного вида геофизических исследований для моделирования рудных тел.

Методическая основа (рис. 1) предусматривает последовательность выполнения следующих этапов:

1. Постановку геологической задачи и выбор объекта моделирования.

2. Анализ имеющихся по объекту геолого- геофизических данных в едином формате и оценка их качества.

3. Выполнение обработки и интерпретации качественных данных.

4. Создание геологического каркаса с геометрией в цифровом формате.

5. Выбор оптимальной единой сетки для комплекса с кубом геофизических данных, материалов по геохимии и геологическому строению объекта (формат сетки GID).

6. Анализ изменения физических свойств пород в зонах с выбранным атрибутом (типом пород).

7. Задание каждому геологическому блоку (каждой ячейке) своего уникального значения с набором параметров.

8. Обучение нейронной сети по классификации типа пород и наличия повышенных содержаний рудных компонентов в зависимости от измеренных физических свойств.

9. Определение веса физических параметров при решении задачи классификации.

10. Вывод о геологической эффективности методов и вероятностных характеристиках прогнозной модели.

Реализация в виде программного обеспечения

Для реализации описанного выше метода разрабатывается программное обеспечение Anthill, включающее построение каркаса геологической модели в формате GID-сетки, где используемые материалы записаны в каталог свойств, характерных для конкретных геологических пород.

Программное обеспечение разработано коллективом инновационной IT-компании, являющейся резидентом новосибирского Академпарка. Оно представлено на международных отраслевых выставках и конференциях, имеет положительные отзывы от образовательных и промышленных организаций. Результаты моделирования верифицировались на данных геофизических измерений, выполненных в Магаданской области и Монголии.

В данной статье показано применение программного обеспечения на Верх-Чемском участке Улантовской перспективной площади, где, проанализировав данные с помощью ПО Anthill, пришли к выводу, что выделенные интрузивные образования можно разделить на два основных класса с разным генезисом. Для подтверждения данного предположения построена основная дискриминационная диаграмма Дж.Пирса Rb-(Y+Nb), где точки составов исследуемых гранитоидов располагаются также в двух зонах. Основная часть их относится к гранитам вулканических дуг, второй ряд проб находится в поле внутриплитных гранитов.

Они отличаются также и содержанием полезных компонентов. Пробы, относящиеся к зоне WPG, содержат более чем на порядок больше меди. Различна и глубина их образования, которая, по данным Rb-Sr метода, составляет 20–30 км для

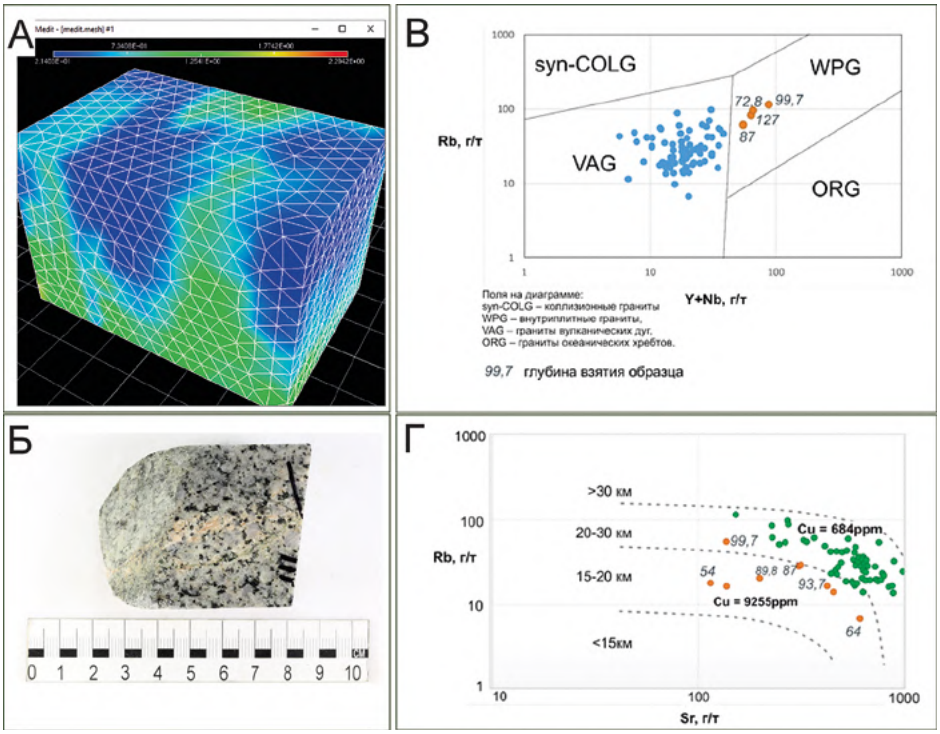


Рис. 2. А) расчетный модуль сетки свойств в ПО AntHill; б) образец породы с базы данных; в-г) диаграммы Rb-(Y+Nb) и Rb(Sr), демонстрирующие разный генезис интрузивных пород Верх-Чемского участка.

основной части интрузива и 15–20 км для гранитоидов зоны WPG.

Заключение

Совместное геолого-геофизическое моделирование данных имеет решающее значение для качественного прогноза содержания полезных компонентов. Чтобы устранить недостатки традиционных подходов в моделировании, таких как влияние боковых объектов и неточность в знании физических параметров пород, при построении модели в ПО Anthill используется трехмерная сеточная модель в сочетании с переносом с эталонного объекта набора свойств. После обучения нейронной сети на данных хорошо изученных участков выполняется построение априорных моделей для новых работ. Показатели прогнозирования позволяют подобрать оптимальные методы исследований, снижающие геологические неопределенности до начала буровых работ, а следовательно, и до несения основных затрат.

Применение к показанному в статье Верх-Чемскому участку данного программного обеспечения позволяет сделать выводы о наличии на данной территории нового рудоносного интрузивного комплекса, представленного порфировидными плагиорганитами. Ранее порфировидные плагиограниты упоминались в отчетах Кадоркина Е. М. Автор выделял их по внешнему виду, с формулировкой: «светлоокрашенные, иногда слабо розоватые породы. Возраст комплекса точно не определен». Их рудоносность ранее установлена не была. Применение же современных технологий и программного комплекса позволило сделать выводы о разном магматическом глубинном источнике, а также доказать почти на порядок превышающее содержание полезных компонентов (меди и молибдена) в порфировидных плагиогранитах (рис. 2).

Литература

1. Jiashun Yao, Lluis Guasch, Michael Warner Neural networks as a tool for domain translation of geophysical data GEOPHYSICS, Volume 88, Issue 3.

2. Е.М.. Кадоркин Отчет партии № 49 о результатах аэропоисковых работ в северной части Салаирского кряжа в 1963 г., г. Новосибирск, 1964 г.

ПРОБЛЕМЫ ОБЕСПЕЧЕНИЯ ПРАВОВОЙ ЗАЩИТЫ ПЕРСОНАЛЬНЫХ ДАННЫХ

ПРИ ИСПОЛЬЗОВАНИИ ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В МЕДИЦИНСКОЙ ДЕЯТЕЛЬНОСТИ

Федеральное государственное бюджетное учреждение науки Институт философии и права Сибирского отделения Российской академии наук
630090, г. Новосибирск, ул. Николаева, 8

Чернусь Надежда Юльевна
кандидат юридических наук, старший научный сотрудник отдела социальных и правовых исследований
тел.: +7-913-903-87-97, preiudicia@yandex.ru

В статье рассматриваются проблемы защиты персональных данных при использовании технологии искусственного интеллекта в медицинской деятельности. Появление машинного обучения, а затем и глубокого обучения в медицинской деятельности ведет к росту уровня и качества диагностики, эффективности организации рабочего процесса и клинических операций, облегчает мониторинг течения заболеваний и их лечения, а также увеличивает эффективность применяемых процедур и в целом результаты лечения пациентов. Тем не менее внедрение и использование технологий искусственного интеллекта связано с определенными информационными и этико-правовыми рисками. Формулируется вывод о наличии в действующем законодательстве разных требований к использованию персональных данных, в том числе в сфере медицинской деятельности. Различные условия обработки и использования предусмотрены в отношении персональных данных, обезличенных персональных данных и анонимных персональных данных.

Концепция использования электронных вычислительных машин для имитации разумного поведения и критического мышления была впервые описана Аланом Тьюрингом в 1950 году в работе «Вычислительные машины и разум» [1], где он описал простой тест, который позже стал известен как «тест Тьюринга», чтобы определить, способны ли компьютеры к человеческому интеллекту. С тех пор технологии искусственного интеллекта претерпели значительные изменения. С появлением машинного обучения, а затем и глубокого обучения возможности применения искусственного интеллекта в различных сферах деятельности настолько расширились, что это создало пер-

спективы замены некоторых сфер человеческой деятельности технологиями искусственного интеллекта. В медицинской деятельности такими перспективами являются создание и развитие персонализированной и роботизированной медицины. Прогностические модели используются для диагностики заболеваний, предсказания терапевтического ответа, профилактической медицины будущего. Применение технологий искусственного интеллекта ведет к росту уровня и качества диагностики, эффективности организации рабочего процесса и клинических операций, облегчает мониторинг течения заболеваний и их лечения, а также увеличивает эффективность применяемых процедур и в целом результаты лечения пациентов. На сегодняшний день развитие технологий искусственного интеллекта достигло такого уровня, когда их использование в узкопрофильных областях оказывается перспективнее использования их в более общих сферах. Тем не менее внедрение и использование технологии искусственного интеллекта связано с определенными информационными и этико-правовыми рисками. Рассмотрим некоторые этико-правовые проблемы внедрения технологии искусственного интеллекта в отдельные сферы медицинской деятельности.

Использование технологии искусственного интеллекта при оказании медицинской помощи связано с обработкой персональной информации о пациенте и сведений обо всех медицинских вмешательствах в отношении него. В данном контексте важно обратить внимание на существование в действующем законодательстве разных концепций правового регулирования оборота персональных данных. С одной стороны, существует необходимость защиты конституционных прав и свобод человека и гражданина, к которым относятся неприкосновенность частной жизни гражданина, личная и врачебная тайна, в связи с чем правовое регулирование защиты персональных данных в Российской Федерации обоснованно определяется положениями публичного законодательства, основанного на императивном методе правового регулирования. С другой стороны, правовое регулирование не должно создавать барьеров для развития инноваций, в противном случае права человека и гражданина не могут быть должным образом реализованы, поскольку внедрение сквозных технологий направлено в том числе и на улучшение качества и уровня жизни населения, персонифицированную медицину, оптимизацию оказания медицинской помощи, повышение качества предоставляемых услуг и т.д.

Характеризуя современное законодательство об охране здоровья граждан, следует отметить, что оно, наряду с законодательством о защите персональных данных, не содержит понятные и достаточные положения для защиты персональных данных о здоровье человека. Согласно ч. 1 ст. 92 Федерального закона от 21.11.2011 № 323-ФЗ «Об основах охраны здоровья граждан в Российской Федерации» [2] (далее — Закон об охране здоровья граждан, Закон № 323-ФЗ) в сфере здравоохранения Российской Федерации ведется учет персональных данных лиц, участвующих в осуществлении медицинской деятельности и фармацевтической деятельности; лиц, обучающихся по образовательным программам среднего профессионального и высшего медицинского образования, образовательным программам среднего профессионального и высшего фармацевтического образования; лиц, которым оказывается медицинская помощь, а также лиц, в отношении которых проводятся медицинские экспертизы, медицинские осмотры и медицинские освидетельствования (далее — персонифицированный учет). Соответственно, для осуществления персонифицированного учета операторы информационных систем в сфере здравоохранения получают информацию от органов и организаций государственной, муниципальной и частной систем здравоохранения и иных организаций в рамках информационного взаимодействия в соответствии с Законом об охране здоровья (ч. 2 ст. 92 Закона № 323-ФЗ). При ведении персонифицированного учета должны обеспечиваться конфиденциальность персональных данных лиц, которые участвуют в системе персонифицированного учета с соблюдением требований, установленных законодательством Российской Федерации в области персональных данных.

В настоящее время в целях обеспечения доступа граждан к услугам в сфере здравоохранения в электронной форме, а также взаимодействия информационных систем в сфере здравоохранения уполномоченным федеральным органом исполнительной власти создается, развивается и эксплуатируется единая государственная информационная система в сфере здравоохранения (далее — единая система). Единая система обеспечивает сбор сведений о лицах, обращающихся за медицинской помощью, включающих, помимо персональ-

ных данных, сведения о медицинской документации и информацию о медицинской организации, в которой медицинская документация создана и хранится. Таким образом, персональные данные пациента являются основой персонифицированного учета в информационных системах здравоохранения, прежде всего, в единой системе, они представляют собой важную составляющую формируемых информационными системами больших данных (big data) в сфере здравоохранения. Безопасность персональных данных пациентов во многом зависит от своевременного обновления программных алгоритмов их идентификации и аутентификации. Хранение персональных данных обеспечивается удаленными серверами, при этом специальных правил функционирования и защиты этих серверов не устанавливается. В настоящее время действуют постановление Правительства РФ от 15.10.2021 № 1753 [3], постановление Правительства РФ от 29.06.2018 № 747 [4], положения которых не определяет ответственных лиц за обновление программных алгоритмов, их идентификацию и аутентификацию.

Итак, на основе персональных данных пациентов формируются базы данных, передача, хранение и обработка размещенных там сведений преследуют прежде всего общественно полезные цели. Однако в российском законодательстве преобладает формальный подход к разработке и применению мер защиты персональных данных, что является одной из причин взломов серверов хранения, а также утечки персональных данных.

Клинические данные, собранные информационными автоматизированными системами, могут быть взломаны и использованы в злонамеренных целях, сводящих к минимуму конфиденциальность и безопасность.

В связи с существующей угрозой использования персональных данных с целью совершения правонарушения, для защиты прав субъектов персональных данных, согласно ч. 1 ст. 9 Федерального закона от 27.07.2006 № 152-ФЗ «О персональных данных» (далее — Закон о персональных данных) субъект персональных данных самостоятельно принимает решение о предоставлении его персональных данных и дает согласие на их обработку. Предусмотрены следующие признаки согласия на их обработку: свобода дачи согласия; волевой характер такого согласия; дача согласия в интересах субъекта персональных данных; конкретность содержания согласия; информированность субъекта, дающего согласие, о целях, способах и операторе обработки персональных данных; осознанность дачи согласия; дача согласия в форме, позволяющей подтвердить факт его получения; дача согласия лично или через представителя.

Также, согласно ч. 1 ст. 6 Закона о персональных данных, обработка персональных данных должна осуществляться с соблюдением принципов и правил, предусмотренных Законом о персональных данных. Обработка персональных данных допускается с согласия субъекта персональных данных на обработку его персональных данных. Далее, в ч. 1 ст. 6 Закона о персональных данных перечислены случаи, когда для обработки персональных данных не требуется получение согласия субъекта персональных данных. В частности, это допускается, когда обработка персональных данных необходима для защиты жизни, здоровья или иных жизненно важных интересов субъекта персональных данных, если получение согласия субъекта персональных данных невозможно (подп. 6 ч. 1 ст. 6), а также, когда обработка персональных данных осуществляется в статистических или иных исследовательских целях (за исключением целей продвижения товаров, работ, услуг на рынке, а также политической агитации) при условии обязательного обезличивания персональных данных. Под обезличиванием персональных данных понимаются действия, в результате которых становится невозможным без использования дополнительной информации определить принадлежность персональных данных конкретному субъекту персональных данных (п. 9 ст. 3 Закона о персональных данных).

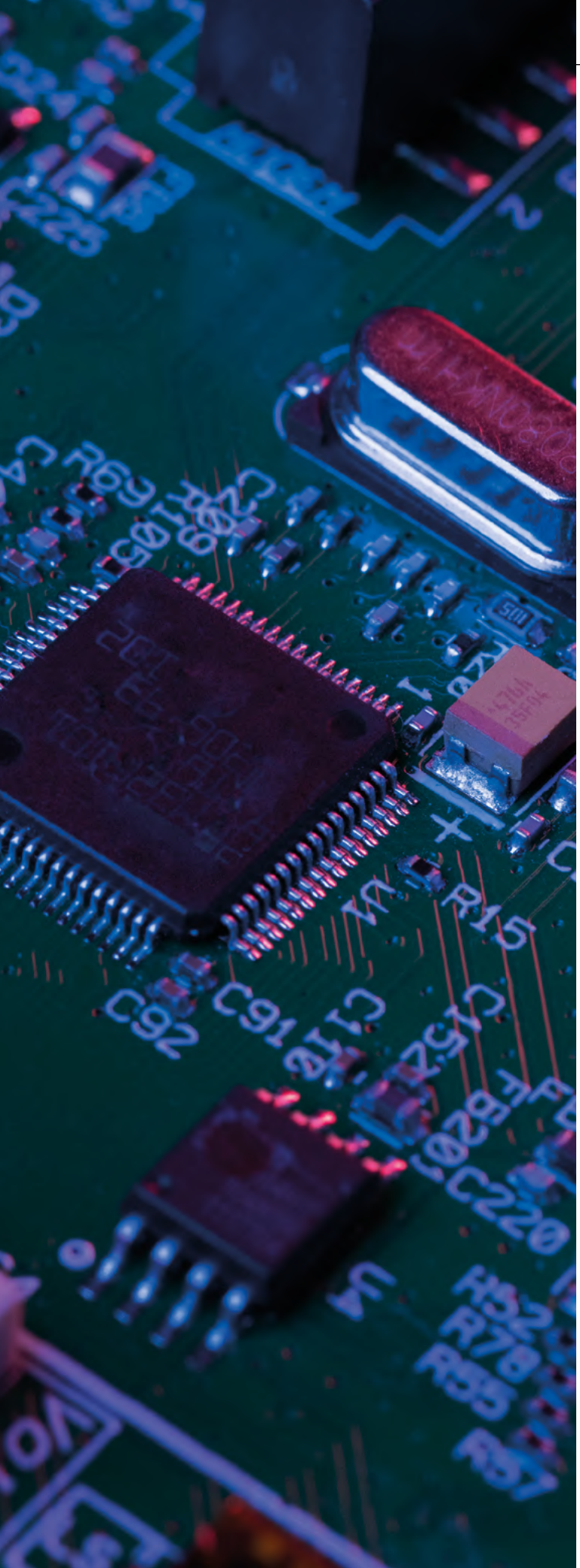
Итак, определенная и весьма значительная часть больших данных (big data), являющихся основой формирования медицинских баз данных, аккумулируется из персональных данных. Как отмечают исследователи, несмотря на то, что в законодательстве определено понятие

обезличивания персональных данных, оно не отражает трансформации защиты данных в условиях развития квантовых технологий и тем самым оказывает негативное воздействие на инновационное развитие. По сути, при возможной идентификации пользователей по дополнительной информации обезличенные данные практически приравнивают к персональным [5]. Вместе с тем обезличенные данные не должны рассматриваться как однородная субстанция. А. И. Савельев отмечает, что обезличенные данные можно условно разделить на те, которые сохраняют связь с физическим лицом и потому могут быть деобезличены, и те, которые необратимо утратили связь с физическим лицом и не могут стать предметом деобезличивания [6]. В первом случае обезличенные данные могут рассматриваться как персональные, но в отношении таких данных возможно установление более льготного режима обработки, в частности за счет расширения круга оснований для их обработки без согласия субъекта. Во втором случае можно говорить о том, что данные перестали быть персональными и могут распространяться без ограничений, установленных законодательством о персональных данных. С точки зрения определенности правового регулирования для обезличенных персональных данных, которые перестали быть персональными, целесообразно ввести особый термин, чтобы избежать терминологической путаницы, которая возникает из-за неоднородности методологий обезличивания и связанной с ней степени обезличенности данных. Такие данные следует отнести к «анонимным данным». Эта категория используется в европейском законодательстве, согласно которому принципы защиты персональных данных не следует применять к анонимной информации, а именно информации, которая не относится к идентифицированному или к поддающемуся идентификации физическому лицу, а также они не должны применяться к персональным данным, анонимизированным таким образом, что субъектов данных вообще или более невозможно идентифицировать (параграф 26 Общего регламента защиты персональных данных (GDPR) Европейского союза [7]).

Таким образом, разные требования по использованию персональных данных, в том числе в сфере медицинской деятельности, предъявляются к персональным данным, обезличенным персональным данным и анонимным персональным данным. Если в отношении персональных данных, в том числе обезличенных, следует применять положения законодательства о защите персональных данных, то в отношении анонимных (анонимизированных) персональных данных применим режим свободного использования, без получения согласия на их обработку и распространение. В частности, при получении персональных данных пациента требуется получение согласия пациента на их обработку, если персональные данные не подлежат обезличиванию; при обезличивании персональных данных пациента, в том числе при анонимизации персональных данных (например, использование изображений, полученных при томографическом, рентгенографическом или ультразвуковом исследовании) получение согласия на обработку персональных данных не требуется. Дальнейшее использование персональных данных (в некоторых случаях обезличенных персональных данных) без получения согласия субъекта персональных данных на их обработку допускается по основаниям, перечисленным в ч. 1 ст. 6 Закона о персональных данных. Использование анонимизированных персональных данных осуществляется без каких-либо ограничений.

Итак, с внедрением во многие сферы человеческой деятельности технологии искусственного интеллекта значение правовой защиты персональных данных возрастает, что отражается на нормативном регулировании, когда право на защиту персональных данных становится самостоятельным субъективным правом наряду с правом на неприкосновенность частной жизни, личной и семейной тайной, защитой своей чести и доброго имени (ч. 1 ст. 23 Конституции РФ). При этом, несмотря на большой потенциал и продвижение искусственного интеллекта в медицине и здравоохранении, его внедрение обусловливает введение новых требований к соблюдению медицинской этики.

Обеспечение баланса положительных и отрицательных последствий внедрения технологии искусственного интеллекта является одним из первоочередных направлений развития законодательства.



Литература

1. Алан Тьюринг. Вычислительные машины и разум. М.: Изд-во АСТ, 2018. 128 с.

2. Официальный интернет-портал правовой информации http://pravo.gov.ru, 22.11.2011

3. Постановление Правительства РФ от 15.10.2021 № 1753 «Об утверждении требований к организационным и техническим условиям осуществления многофункциональными центрами предоставления государственных и муниципальных услуг размещения или обновления в единой системе идентификации и аутентификации сведений, необходимых для регистрации физических лиц в данной системе, размещения биометрических персональных данных в единой информационной системе персональных данных, обеспечивающей обработку, включая сбор и хранение, биометрических персональных данных, их проверку и передачу информации о степени их соответствия предоставленным биометрическим персональным данным физического лица, с использованием программно-технических комплексов» // Официальный интернет-портал правовой информации http://pravo.gov.ru, 18.10.2021.

4. Постановление Правительства РФ от 29.06.2018 № 747 «Об установлении требований к фиксированию действий при размещении в электронной форме в единой системе идентификации и аутентификации сведений, необходимых для регистрации физического лица в указанной системе, и иных сведений, предусмотренных федеральными законами, а также при размещении биометрических персональных данных физического лица в единой информационной системе персональных данных, обеспечивающей обработку, включая сбор и хранение, биометрических персональных данных, их проверку и передачу информации о степени их соответствия предоставленным биометрическим персональным данным физического лица» // Официальный интернет-портал правовой информации http://pravo.gov.ru, 02.07.2018.

5. Лескина Э. И. Квантовые технологии и перспективы правового регулирования защиты данных // Юрист. 2022. № 11. С. 3.

6. Савельев А. И. На пути к концепции регулирования данных в условиях цифровой экономики // Закон. 2019. № 4. С. 180.

7. Общий регламент защиты персональных данных (GDPR) Европейского союза [Электронный ресурс]: Режим доступа: https://gdpr-text.com/ru/ (дата обращения: 11.06.2023).

3

Последние достижения ИТ

92	Staffcop Enterprise
96	Семейство программно-аппаратных комплексов обработки больших потоков данных в режиме реального времени
100	Разработка и верификация автоматных программ для систем управления
107	Методика оценки состояния и выбора структуры высоконадежной распределительной сети на основе применения меры неопределенности информации
111	FIS Platform: low-code платформа для автоматизации финансового сектора

STAFFCOP ENTERPRISE

ООО «АТОМ БЕЗОПАСНОСТЬ»
630128, г. Новосибирск,
ул. Кутателадзе , 4Г

Хаустов Иван Юрьевич
генеральный директор ООО «АТОМ БЕЗОПАСНОСТЬ»
hij@staffcop.ru

Информационная безопасность — один из наиболее актуальных вопросов современного бизнеса. С каждым годом количество утечек конфиденциальной информации и кибератак на государство и бизнес растет, что приводит к значительным финансовым потерям и репутационным проблемам. Компаниям необходимо принимать проактивные меры для защиты своих данных и снижения рисков утечки информации.

Одним из основных инструментов для обеспечения безопасности информации является программное обеспечение, которое позволяет отслеживать потоки информации и системы аналитики, а также предупреждать нежелательные действия. В числе наиболее эффективных продуктов на рынке может быть назван адаптированный под современные реалии отечественный программный продукт для расследования и предотвращения инцидентов внутренней информационной безопасности — Staffcop Enterprise компании ООО «АТОМ БЕЗОПАСНОСТЬ».

Продукт Staffcop разработан с учетом современных требований к безопасности и обладает рядом преимуществ:

- Во-первых, он позволяет отслеживать потоки информации и системы аналитики в режиме реального времени. Таким образом, компания может быстро реагировать на потенциальные угрозы и предупреждать нежелательные действия.
- Во-вторых, данный продукт может работать на удаленном компьютере, не находящемся в локальной сети компании. Это позволяет обеспечить безопасность данных и информации даже в случае удаленной работы сотрудников.
- В третьих, программа обладает функцией автоматического оповещения о возможных инцидентах информационной безопасности. Это позволяет быстро реагировать на угрозы и предотвращать их.

Кроме того, продукт имеет возможность настройки под нужды конкретной компании, что позволяет обеспечить максимальную защиту данных и информации.

Сама идея создания такого продукта возникла в 2004 году на одной из IT-конференций. Так появился прототип программного обеспечения Staffcop Standard. Позже было решено переписать и адаптировать его для корпоративных нужд, что привело к созданию Staffcop Enterprise.

С 2012 года компания активно развивается — получила сертификат ФСТЭК, стала резидентом Сколково и Академпарка. Covid-19 придал новый импульс развитию компании, поскольку резко выросло количество работников «на удалёнке», эффективность работы которых необходимо было отслеживать. Многие компании не были готовы к такому развитию событий с точки зрения наличия нужного программного обеспечения, и продажи Staffcop Enterprise резко выросли. В 2021 году компания приняла решение найти более крупного и технологически зрелого партнера, который был бы заинтересован в развитии продукта, но при этом готов сохранить автономность компании. В результате ООО «АТОМ БЕЗОПАСНОСТЬ» вошло в группу компаний «СКБ КОНТУР». Процесс импортозамещения и уход зарубежных DLP-систем также открыл новые перспективы продукту, особенно в крупных корпорациях госсектора и оборонной промышленности.

К сегодняшнему дню Staffcop отличается:

- хорошим решением задач пользователей внутренней информационной безопасности, сопоставимым с импортными аналогами;
- архитектурой, использующей СПО;
- логикой в настройке процессов, которая позволяет закрывать 100% потребностей клиентов в расследовании инцидентов, связанных с утечкой информации.

Также в апреле 2023 года компания представила новую улучшенную версию Staffcop Enterprise. Какие семь главных изменения были внесены в релиз 5.2 Staffcop Enterprise?

1. Горизонтально масштабируемая архитектура сняла ограничения на количество рабочих станций под контролем.
2. Интеграция с российскими бизнес-коммуникаторами VK Teams, VK Work Mail, eXpress.
3. Изменение принципа работы со скриншотами, что оптимизировало нагрузку на места хранения.
4. Консолидированные отчеты, которые позволяют делать аналитику более глубокого уровня без дополнительной нагрузки на администратора системы и в режиме офлайн.
5. Реализован перехват HTTPs-соединений.
6. Утилита мониторинга рабочих станций, которая обеспечивает непрерывный сбор данных для полной доказательной базы в случае инцидентов.
7. Мониторинг с созданием теневых копий работы CD\DVD-файлов.

Кроме того, организация обладает импортонезависимостью, входит в реестр отечественного ПО.

Разработка программного обеспечения для защиты информации даст множество преимуществ:

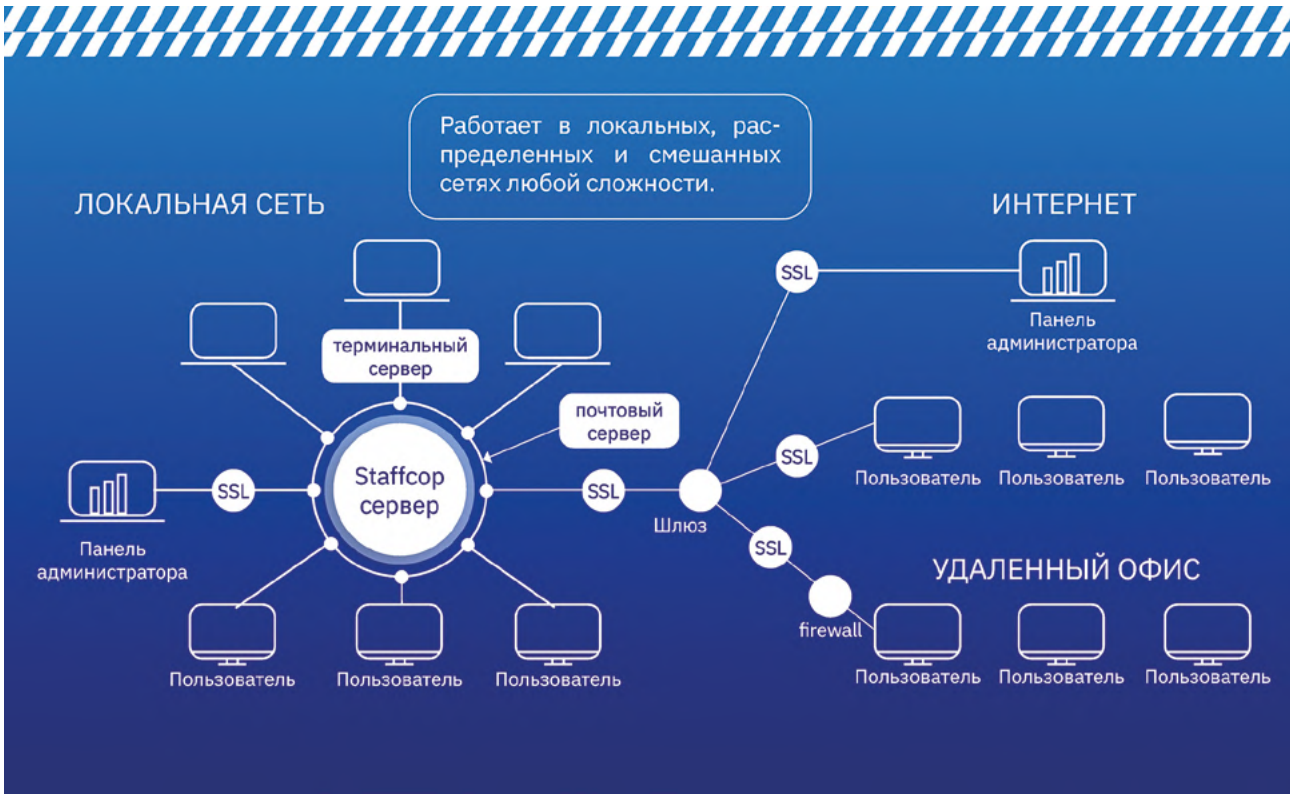


Рис. 1. Архитектура Staffcop.

1. Безопасное хранение конфиденциальных данных: такое программное обеспечение помогает сохранить конфиденциальность и безопасность любой информации, хранящейся на компьютерах и серверах.

2. Защита бизнеса: для успешной работы любого бизнеса важно сохранить свои конфиденциальные данные и информацию о покупателях/партнерах, и конфиденциальность этих данных должна быть гарантирована. Разработка программного обеспечения для защиты информации поможет предотвратить утечки данных, а также собрать легитимную доказательную базу в случае, если инцидент все-таки свершился.

3. Укрепление национальной безопасности: жизненно важно иметь программное обеспечение, которое может защитить государственные организации и критическую инфраструктуру от киберугроз и промышленного шпионажа.

4. Защита личных данных: с развитием цифровых технологий личные данные стали более уязвимыми, вся наша жизнь происходит в цифровом пространстве. Программное обеспечение для защиты информации поможет защитить личные данные пользователей.

Таким образом, разработка программного обеспечения для защиты информации является важной для обеспечения безопасности в различных областях, на всех уровнях:

1. Государственный сектор — правительственные агентства, военные учреждения, полиция.

2. Финансовая отрасль — банки, страховые компании, инвестиционные фонды.

3. Розничная торговля — магазины, супермаркеты, интернет-магазины.

4. Здравоохранение — больницы, клиники, лаборатории.

5. Транспортная отрасль — авиакомпании, железнодорожные операторы, автоперевозчики и пр.

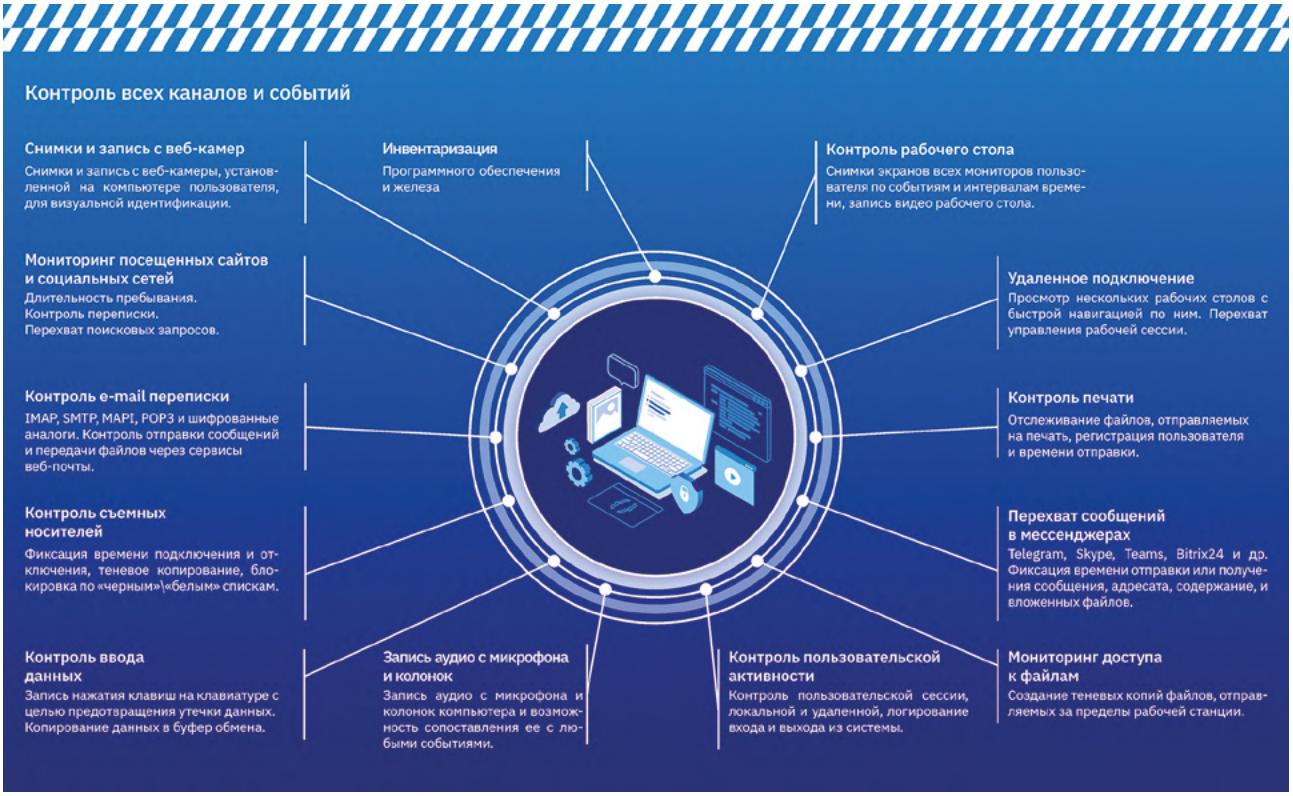


Рис. 2. Основные функции продукта Staffcop.

Хаустов Иван Юрьевич

Цитата генерального директора:

«Компания «АТОМ БЕЗОПАСНОСТЬ» является одним из лидеров в области разработки программных решений для обеспечения информационной безопасности. Наша компания разрабатывает Staffcop — инновационное решение для контроля информационных потоков и расследования инцидентов внутренней безопасности. Программное обеспечение позволяет удостовериться в том, что пользователи не используют конфиденциальную информацию неправомерно, а также предупреждать о любых попытках утечки данных».

ООО «АТОМ БЕЗОПАСНОСТЬ» всегда ищет новых специалистов, которые могут принести новые идеи и помочь в разработке новых программных решений в области обеспечения информационной безопасности.

Основной язык разработки продукта — C++. Сервер пишется на Python под Django. Основной стек технологий: Python, Pytest, Selenium WebDriver, PyWinAuto, Allure, Teamcity, VMware vSphere, Redmine, Testlink.

СЕМЕЙСТВО ПРОГРАММНО-АППАРАТНЫХ КОМПЛЕКСОВ

ОБРАБОТКИ БОЛЬШИХ ПОТОКОВ ДАННЫХ В РЕЖИМЕ РЕАЛЬНОГО ВРЕМЕНИ

Общество с ограниченной ответственностью «СофтЛаб-НСК»
630090, г. Новосибирск,
пр. Академика Коптюга 1а, офис 311

Травина Ирина Аманжоловна
генеральный директор

Лаврентьев Михаил Михайлович
доктор физико-математических наук, профессор, руководитель проектов

Шадрин Михаил Юрьевич
начальник отдела мультимедиа

Таранцев Игорь Геннадьевич
начальник отдела разработки программного обеспечения

Лысаков Константин Фёдорович
кандидат технических наук, начальник отдела аппаратных разработок
тел.: +7-913-908-18-26

Компания СофтЛаб-НСК с 1991 года работает в области обработки мультимедиа-данных в темпе поступления. Телевидение является основной областью потребления разрабатываемых продуктов, обеспечивающих долговременную надежную работу разрабатываемых решений для приема, обработки и передачи видеопотоков. Также продукция сегодня применяется для организации вещания и трансляции спортивных событий, обеспечивая возможность коммутации множества источников видео, оперативного переключения, микширования, а также функционирования системы замедленных повторов, что актуально для видеосудейства.

История развития продуктов

В 1993 году была разработана плата расширения в персональный компьютер IBM-PC, который можно было свободно купить и который уже имели компании, работающие в области 3D-графики. Программное обеспечение делало возможным воспроизведение синтезированного фрагмента ролика длиной до нескольких секунд в виде стандартного телевизионного сигнала (ПАЛ, СЕКАМ, Y/C, YUV или RGB). В 1996 году вышла вторая плата — «Форвард», на которой использовался аппаратный кодек JPEG, что позволило на порядок понизить требования к передаваемым потокам данных и обеспечило непрерывный режим работы системы. Решения на платах «Форвард» могли воспроизводить ролики непрерывно неограниченное время в режиме 24/7/365–24 часа в сутки, 7 дней в неделю, 365 дней в году (раз в год нужно выключить и обслужить ПК, чтобы он надежно работал весь следующий год). Решения «Форвард» базировались на операционной системе Windows. В те времена именно на Windows работали большинство компаний, занимающихся компьютерной графикой. Топовые системы компании Silicon Graphics были очень дороги для российских потребителей, Apple пришел в видеомонтаж слишком поздно. Системы на базе UNIX всегда были слишком сложными для творческих работников (художников, дизайнеров). На российском рынке компьютерной графики тогда использовался 3DStudio Max под Windows. Операционная система Windows не является системой реального времени. Задержка выполнения любой программы может достигать многих сотен миллисекунд или даже секунд. Тем не менее для пользователя эти задержки являются приемлемыми, и пришлось разрабатывать решение именно под Windows.

Телевизионный сигнал был жестко таймирован — система обязана каждые 40 миллисекунд генерировать очередное изображение. Эта задача возлагалась на плату ввода-вывода, которая использует внутреннюю память для буферизации очередного изображения и занимается формированием телевизионного сигнала в соответствии с требованиями стандарта. По DMA (Direct Memory Access) плата самостоятельно копирует данные из памяти компьютера во внутренний буфер и, наоборот, из внутреннего буфера в память компьютера (входные и выходные потоки видео). Работа обычных программ на компьютере не может помешать работе DMA, поэтому данные успевают скопироваться за время кадра. Также плата каждые 40 миллисекунд инициирует через аппаратные прерывания вызов программного обеспечения — драйвера платы. Хотя обработка прерывания может задерживаться на какое-то время, практически всегда оно не превышает десятка миллисекунд. Получив управление, драйвер передает в плату информацию о том, откуда нужно копировать следующий блок данных.

Плата обеспечивает буферизацию данных на уровне нескольких кадров — пока плата выводит на выход один кадр и копирует внутрь себя следующий кадр, драйвер успевает дать ей задание еще на кадр вперед. Однако из опыта эксплуатации известно, что операционная система может «зависнуть» на десятки кадров, вплоть до секунды. Поэтому следующий уровень буферизации реализован на программном уровне. Драйвер платы выполняет роль микшера/коммутатора. По запросу программы он отводит в оперативной памяти компьютера большой буфер (на несколько десятков кадров) и организует на этом буфере очередь. Программа генерирует последовательность изображений и записывает их в очередь, следя за тем, чтобы она была постоянно заполнена. Драйвер в прерывании удаляет из очереди уже показанные изображения и прописывает в плату информацию о следующем изображении. Как правило, приложение создает пару очередей — одна используется для воспроизведения текущего списка роликов, а вторая заполнена первыми кадрами альтернативного списка роликов. В любой момент времени оператор может нажать кнопу старта альтернативного списка, тогда программа даст команду драйверу, и он переключится с одного буфера на другой. Поскольку данные в буфере уже заполнены, переключение происходит максимально быстро — буквально в следующем кадре. Задержка переключения соответствует привычным для пользователя задержкам и не превышает времени реакции человека. Таким

Рис. 1. Решение для визуального контроля процесса тренировок космонавтов в НИИ ЦПК имени Ю.А. Гагарина.

образом, аппаратная буферизация гарантирует стабильность выдачи данных на уровне десятков миллисекунд, а программная буферизация в оперативной памяти компьютера — стабильность работы на несколько секунд вперед. Эта технология успешно работает и сейчас.

За 30 лет сотрудниками компании было создано семейство решений, ориентированных на обработку больших потоков данных в темпе поступления этих данных. Для большинства из них технология буферизации данных на очередях в оперативной памяти компьютера является ключевой:

- TitleEngine — технология генерации и совмещения многослойных полупрозрачных титров, позволяющая совмещать гарантированную плавность движения с интерактивностью управления (в любой момент можно включить или выключить любой слой титров).

- PostPlay — технология буферизованной записи сжатых видео- и звуковых данных на жесткие диски персонального компьютера, позволяющая минимизировать нагрузку на диски и получить максимальную скорость чтения/записи данных на диск при многоканальной записи/чтении.

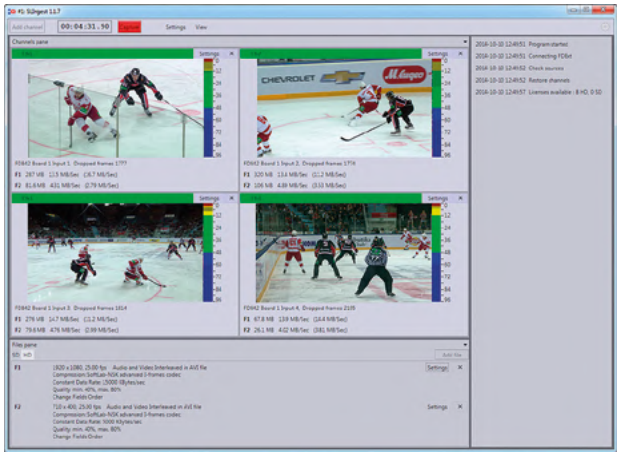
- AutoDetect — технология анализа потока видео/звука для детектирования в нем записанных ранее фрагментов видео/звука. При совпадении входного видео/звука с одним из фрагментов система генерирует управляющее событие, которое может, например, инициировать с точностью до кадра воспроизведение списка роликов.

Разработанное для удобного воспроизведения видеоповторов программно-аппаратное решение также нашло применение в ФГБУ «Научно-исследовательский испытательный центр подготовки космонавтов имени Ю. А. Гагарина».

Проблема надежности

Одной из проблем создания сложных программно-аппаратных систем для обработки больших потоков данных является проблема надежности, поскольку с усложнением систем обработки приходится использовать программное обеспечение (библиотеки) от сторонних производителей. Если программный код собственной разработки можно отладить и быть уверенным в его высокой надежности, то внешние библиотеки выступают в роли «черного ящика». К сожалению, такие технологии Windows, как DirectShow, славятся своей нестабильностью.

❷



❸

В рамках проделанных исследований было разработано следующее решение: если время отказа не должно превышать 100 миллисекунд, то в решениях компании рекомендуется использовать горячее резервирование и специальное устройство watch dog. Горячее резервирование означает полное дублирование системы вторым компьютером с идентичным набором «железа» и софта. Выходы обоих систем поступают на коммутатор резерва, который управляется устройством watch dog. Логика управления устройством watch dog следующая. Пользовательское приложение регулярно сообщает драйверу, что у него всё будет хорошо еще некоторое время. Драйвер в каждом прерывании сравнивает это время с текущим и оценивает глубину заполненности очереди на воспроизведение. Если данные в очереди есть и приложение недавно рапортовало о хорошем статусе, то драйвер посылает сигнал на

4

устройство watch dog, что все хорошо. Если в течение пары кадров на устройство watch dog не приходил сигнал, то оно автоматически переходит в режим «авария» и переключает коммутатор резерва. Традиционно устройство watch dog подключают к основному компьютеру, поэтому переключение происходит при любом отказе — при отключении питания, при отказе платы ввода-вывода, при «зависании» компьютера, при «зависании» пользовательских программ. Для упрощения работы оператора и уменьшения числа человеческих ошибок пользовательские приложения компании «СофтЛаб-НСК» поддерживают режим «зеркалирования», когда действия оператора в одном приложении автоматически пересылаются в приложение на зеркалируемом компьютере, то есть оператору не нужно повторять действия на обоих.

Если же время отказа может быть достаточно большим, в пределах десятков секунд или минут, то в решениях компании достаточно использовать разделение системы на отдельные программные модули, исполняющиеся в независимых процессах. При «взрыве» любого процесса специальная система автоматически запускает новый процесс вместо выбывшего из работы. Такая модульность системы на уровне отдельных процессов, конечно же, порождает дополнительные трудности, связанные не только с управлением, но и с передачей больших потоков данных между процессами. Для этого опять же необходимо использовать буферизацию данных в оперативной памяти компьютера. Причем решения компании избегают лишних копирований данных — модуль «поставщик данных» заполняет изображение в памяти буфера, а модуль «приемник данных» берет данные для анализа прямо из памяти буфера.

Также разделение на процессы успешно решает проблему надежности работы системы в целом при непродолжительной перегрузке. Задавая разные приоритеты разным программным модулям, можно управлять временем работы каждого из них. При полной загрузке процессора в первую очередь будут работать только важные модули с высочайшим приоритетом. К сожалению, в обычных операционных системах нет аналогичного управления приоритетом выполнения файловых операций. Поэтому в решениях компании используется собственный модуль приоритизации запросов к файловой системе. Это решает проблему при следовании рекомендованному рабочему процессу, но если пользователь запустит стороннее приложение, то система в целом может не успеть своевременно отработать все запросы на чтение/запись. Технология PostPlay как раз гарантирует своевременность записи и чтения данных.

Выводы

В настоящее время компания «СофтЛаб-НСК» выпускает широкий спектр плат ввода-вывода, включая платы для работы с телевизионным сигналом с разрешением UHD4K (3840x2160 60 к/с). В современных платах всё больший акцент делается на программном обеспечении — в них не используется аппаратных кодеров видео, поскольку вся обработка, включая кодирование и декодирование видео, выполняется стандартными средствами современного персонального компьютера. Решения компании мигрируют на операционную систему Linux, в первую очередь это касается решений для современного цифрового телевидения. Увеличение потоков данных, передаваемых в современном телевидении, ставит новые задачи, которые они успешно решают.

2. Пример виртуальной студии, работающей с платами компании «СофтЛаб-НСК» для захвата «живого» актера.

3. Решение для многоканальной записи с автоматической синхронизацией данных.

4. Линейка современных плат компании «СофтЛаб-НСК».

РАЗРАБОТКА И ВЕРИФИКАЦИЯ АВТОМАТНЫХ ПРОГРАММ ДЛЯ СИСТЕМ УПРАВЛЕНИЯ

Институт систем информатики им. А.П. Ершова СО РАН

Федеральное государственное бюджетное учреждение науки Институт систем информатики им. А.П. Ершова Сибирского отделения Российской академии наук
630090, г. Новосибирск, пр. Академика Лаврентьева, 6

Методы автоматного программирования иллюстрируются на примере нетривиальной задачи управления движением автомобилей на мосту.

Шелехов Владимир Иванович
заведующий лабораторией системного программирования, кандидат технических наук, доцент
тел.: (383) 330-27-21, vshel@iis.nsk.su

Тумуров Эрдэм Гармаевич
младший научный сотрудник
тел.: (383) 330-27-21, erdemus@gmail.com

Шелехов Владимир Иванович, заведующий лабораторией системного программирования, кандидат технических наук, доцент

Тумуров Эрдэм Гармаевич, младший научный сотрудник

Введение

Последние 30 лет характеризуются кризисом базисных технологий программирования при бурном росте информационных технологий в самых различных областях приложений. Технология разработки программ для систем управления — одна из базисных. В мире существуют сотни таких технологий, не менее четырех в новосибирском Академгородке. Event-B [1] — одна из наиболее популярных технологий. Это метод формальной спецификации и верификации систем в программной и системной инженерии, успешно используемый при разработке производственных систем управления, особенно в железнодорожном транспорте и метрополитене. Сила метода Event-B в том, что он позволяет обнаруживать многие серьезные ошибки с помощью доказательства теорем, что нереализуемо другими известными методами.

Задача управления движением автомобилей на мосту является показательным примером в руководстве по Event-B [1]. На этой задаче детально описывается технология Event-B, существенной частью которой является весьма интересный, но малоизвестный у нас, метод уточнения (refinement): решение задачи конструируется

в виде цепочки моделей от простой к более сложной. Метод уточнения является универсальным, но его применение для конкретной задачи не всегда является полезным, в частности для данной задачи.

Критерий полезности метода в конкретном применении сформулирован в работе [2]. В частности, полезность метода объектно-ориентированного программирования определяется возможностью инкапсуляции (упрятывания) разнообразных деталей реализации, предоставляя наружу более простой интерфейс. Если же не происходит существенного упрощения внутреннего интерфейса программы, то объектно ориентированные конструкции просто загромождают программу, удлиняя и усложняя ее. К сожалению, ошибки такого рода совершают многие, в том числе специалисты высшего уровня квалификации.

Более простое и короткое решение задачи управления движением на мосту было представлено в нашей работе [2] с использованием технологии автоматного программирования. При верификации данного решения в системе Event-B было найдено четыре нетривиальных ошибки. Сам факт обнаружения этих ошибок, требующий квалификации инженера-верификатора, показывает, что и наше решение нетривиально. Дальнейший анализ выявил другие недостатки. Таким образом, определился новый вызов к разработке адекватных методов решения подобных задач.

Задача управления движением автомобилей на мосту

Имеется мост, соединяющий материк и остров. Мост узкий и позволяет двигаться автомобилям по нему только в одну сторону. Имеются два светофора, установленных при въезде на мост с материка и с острова. У каждого светофора два цвета: красный и зеленый. Автомобилям запрещено движение на красный свет при въезде на мост. Имеются четыре сенсора. Каждый сенсор находится на некотором участке автомобильной трассы и способен фиксировать факт нахождения автомобиля на нем. Первый сенсор находится перед въездом на мост с материка. Второй сенсор — на мосту перед съездом на остров. Третий сенсор — перед въездом на мост с острова. Четвертый сенсор стоит на мосту перед съездом на материк.

Число автомобилей, которые могут находиться на острове, ограничено. Необходимо построить

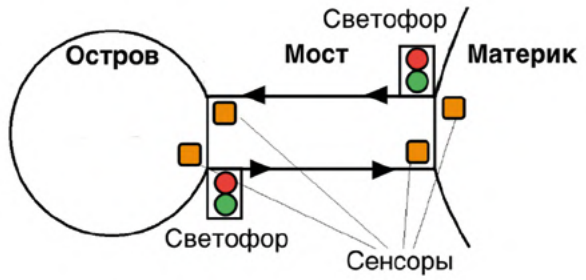


Рис. 1. Схема моста со светофорами и сенсорами.

программу, которая переключает светофоры, используя показания сенсоров, и таким образом обеспечивает безопасное движение автомобилей по мосту.

Автоматное программирование

Автоматное программирование [2–4] ориентировано на класс реактивных систем (в частности, систем управления), реализующих взаимодействие с внешним окружением системы и реагирующих на определенный набор событий (сообщений). Реактивная система состоит из набора параллельных взаимодействующих процессов, работающих на некотором общем поле переменных, называемом состоянием системы. Каждый процесс представлен независимой автоматной программой, определяемой в виде автоматной композиции операторов (сегментов).

Автоматное программирование реализуемо на любом языке программирования. Автоматная композиция может быть закодирована другими конструкциями языка. Здесь в качестве базисного языка используется подмножество языка Event-B [1].

Сегмент должен быть атомарным: исполнение сегмента должно быть единым, неделимым актом; это означает, что до завершения исполнения сегмента все другие параллельно исполняемые процессы приостанавливаются.

Задача разработки реактивной системы формулируется в виде набора требований. Требование— утверждение, определяющее потребность и связанные с ней измеримые условия и ограничения. Требования к реактивной системе включают требования окружения и функциональные требования, определяющие поведение системы. Существенными являются также

нефункциональные требования надежности, безопасности, защищенности, отсутствия дедлоков и другие. Разработка требований должна проводиться в соответствии со стандартом ISO/ IEC/ IEEE29148 [5].

Достаточно часто автоматная программа строится простым переписыванием функциональных требований. Однако многие распределенные системы со сложными взаимосвязями, в частности, системы управления роботами и БПЛА, требуют особых методов разработки. Некоторые из таких методов представлены здесь на задаче управления движением на мосту.

Универсальный контроллер для сенсора

Решение задачи в работе [2] было предложено для исправления недостатков исходного решения в руководстве по Event-B [1]. Функционирование сенсоров реализуется независимыми процессами. Они должны подсчитывать число автомобилей на мосту и на острове. А управление светофорами было решено реализовать пятым головным независимым процессом. Однако и наше решение оказалось сложным, провоцирующим нетривиальные ошибки. Необходимость пятого процесса не была обоснована. В действительности он появился как результат переработки головного процесса исходного решения [1].

Необходим метод адекватной декомпозиции решения подобных задач. **Первый метод**, предлагаемый здесь, — реализация снизу, начиная от простейших компонент. Сначала разрабатываются программы для сенсоров, и лишь затем на их базе выстраивается полная программа. Программы контроллеров для четырех сенсоров разные. Однако у них есть общая часть. На первом этапе построим универсальный контроллер для всех сенсоров. На втором этапе проводится специализация универсального контроллера для каждого из четырех сенсоров.

Построение универсального контроллера — это **второй метод** разработки общей модели для нескольких программ. Он проще метода уточнений для Event-B [1] и более универсален.

Состояние универсального контроллера определим следующей секцией:

```
section {
    on: BOOL:= false
    off: BOOL:= false
}
```

Переменной on присваивается true в тот момент, когда автомобиль заедет на сенсор. Переменной off присваивается true сразу после того, как очередной автомобиль съедет с сенсора. Данные действия реализуются сенсором. Будем считать, что сенсор в соответствующий момент запускает один из процессов:

```
process On {on:= true}
process Off {off:= true}
```

Отметим, что сенсор всего лишь присваивает true переменным on и off, но никак не реагирует на их значения.

Универсальный контроллер определяется процессом Sensor, программа которого представлена автоматной композицией из двух сегментов, запускаемых для обработки событий включения и выключения сенсора:

```
process Sensor {
    1: inv not off; on → on:= false; Act1; #2
    2: inv not on; off → off:= false; Act2; #1
}
```

Находясь в управляющем состоянии “1”, контроллер ожидает события включения сенсора, реализуемого срабатыванием оператора on:= true. Когда событие случается (автомобиль заезжает на сенсор), гасится признак события on присваиванием on:= false и исполняется обработчик события — оператор Act1. Далее контроллер переходит в управляющее состояние “2”, ожидая события выключения сенсора (автомобиль съезжает с сенсора). При выключении сенсора гасится признак off, выполняется обработчик Act2 и контроллер переходит в управляющее состояние “1”.

Управляющие состояния снабжаются инвариантами not off и not on, которые должны быть истинны, когда контроллер находится в начале соответствующего сегмента.

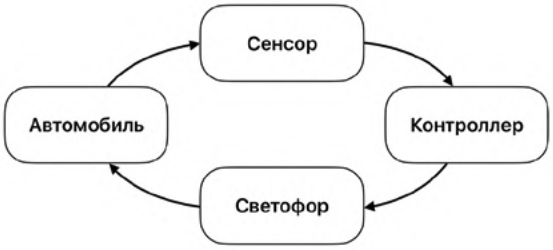


Рис. 2 Взаимодействие компонентов системы управления движением на мосту.

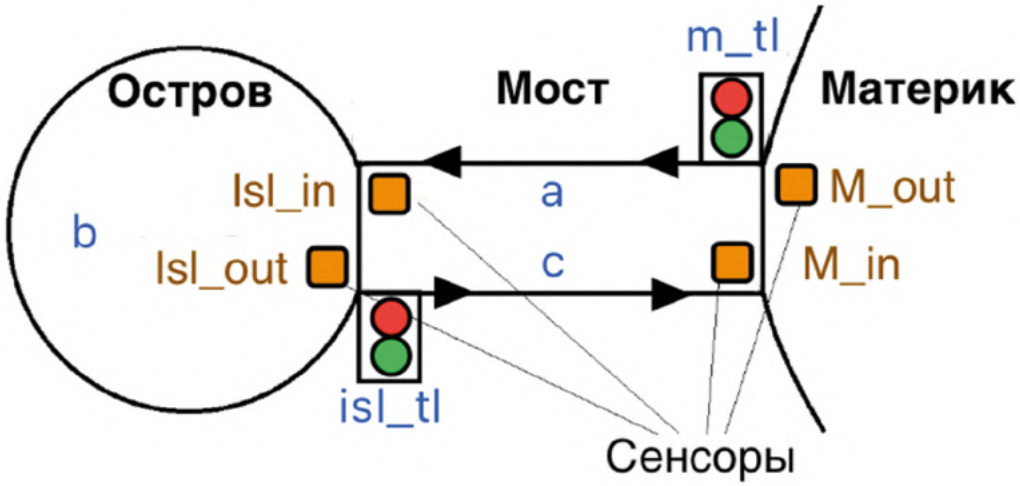


Рис. 3. Состояние программы управления движением на мосту.

Инварианты не исполняются. Они используются для верификации.

Семантика оператора "→" определяется следующей эквивалентной программой контроллера.

```
process Sensor {
    1: inv not off; if (on) {on:= false; Act1; #2} else #1
    2: inv not on; if (off) {off:= false; Act2; #1} else #2
}
```

Автоматную композицию можно закодировать другими конструкциями языка программирования. В языках Си и C++ автоматная композиция процесса Sensor кодируется следующим образом:

```
void Sensor() {
    M1: if (on) {on = false; Act1; goto M2} else goto M1
    M2: if (off) {off = false; Act2; goto M1} else goto M2
}
```

В языках без оператора перехода применяется оператор выбора switch по управляющим состояниям, представленным значениями типа перечисления:

```
void Sensor() {
    type State = enum {1, 2};
    State state = 1;
    while (true) {
        switch (state) {
            case 1: if (on) {on = false; Act1; state=2} else
            break
            case 2: if (off) {off = false; Act2; state=1}
 }}}
```

Контроллер должен успевать отслеживать изменения показаний сенсора. Это следует сформулировать в виде следующих требований.

RFS1: Включение сенсора при появлении на нем автомобиля должно быть зафиксировано и обработано раньше, чем другой автомобиль заедет на сенсор.

RFS2: Выключение сенсора при съезде с него автомобиля должно быть зафиксировано и обработано раньше, чем другой автомобиль съедет с этого сенсора.

Построение контроллеров для четырех сенсоров

Для написания программы контроллера для каждого из четырех сенсоров достаточно построить программы обработчиков Act1 и Act2 для событий включения и выключения сенсора в составе программы универсального контроллера. Назначение контроллеров — определять число автомобилей на мосту и на острове по показаниям сенсоров. На основе этих данных далее реализуется управление светофорами.

Состояние программы управления движением на мосту определяется значениями следующих переменных (см. рис. 3):

a — число автомобилей на мосту, едущих с материка на остров;

b — число автомобилей на острове;

c — число автомобилей на мосту, едущих с острова на материк;

m_tl — цвет светофора на материке (зеленый или красный);

isl_tl — цвет светофора на острове (зеленый или красный).

Состояние программы представлено следующей секцией:

```
section St0 {
    const d: NAT (*максимальное число автомо
                    билей на острове*)
    axiom d>0
    a: NAT:= 0
    b: NAT:= 0
    c: NAT:= 0
    type Colour = enum {red, green} (*тип цвета
                                    светофора*)
    m_tl: Colour:= green (*сначала зеленый*)
    isl_tl: Colour:= red (*сначала красный*)
    inv a + b <= d (*инвариант: число автомоби
                    лей на острове <= d *)
    inv m_tl = red ∨ isl_tl = red (*оба светофора
                        не могут быть зелеными*)
}
```

Максимально допустимое число автомобилей на острове определено константой d, причем d>0; при d=0 будет дедлок. В секции определены значения переменных в начальный момент работы программы управления движением на мосту. Инвариант a + b <= d фиксирует ограничение по числу автомобилей на острове, учитывая, что автомобили на мосту беспрепятственно попадут на остров. Инвариант m_tl = red ∨ isl_tl = red отражает требование безопасности. В решении [2] вместо c используется переменная a. Такая модификация уменьшает число переменных в состоянии, однако усложняет семантику переменной a, что спровоцировало ошибку.

Третий метод: автономная разработка программы для каждой компоненты, исходя только из ее логики, и пока без детального отслеживания возможных взаимодействий между разными компонентами. Конфликты между компонентами будут разрешены на следующем этапе разработки.

Начнем с построения более простых контроллеров для съезда с моста, модифицирующих значения переменных a, b и c.

```
process M_in {
    1: on → on:= false; #2
    2: off → off:= false; c:=c-1; #1
}
process Isl_in {
    3: on → on:= false; #4
    4: off → off:= false; a:=a-1; b:=b+1; #3
}
```

Эти контроллеры просты. Реакция на событие on в них отсутствует. Поэтому сегменты “1” и “3” не нужны и в дальнейшем будут удалены.

Контроллер Isl_out для въезда на мост с острова в случае события on сначала проверяет показание светофора isl_tl на острове. Если светофор зеленый, то реализуется переход на сегмент “7”, где происходит ожидание события off, после чего модифицируюся значения переменных b и c. Если светофор isl_tl красный, происходит переход на сегмент “6”, где ожидается событие a=0, сигнализирующее, что мост пуст. В случае a=0 происходит изменение цвета двух светофоров и переход на сегмент “7”.

```
process Isl_out {
    5: on → on:= false; if (isl_tl=green) #7 else #6
    6: a=0 → isl_tl:=green; m_tl:=red; #7
    7: off → off:= false; b:=b-1; c:=c+1; #5
}
```

Контроллер M_out для въезда на мост с материка конструируется аналогично контроллеру Isl_out с заменой светофора isl_tl на m_tl и условия a=0 на c=0. Однако контроллер M_out сложнее, потому что здесь дополнительно нужно обеспечить ограничение a + b <= d. Допустим, очередной автомобиль заехал на сенсор, т.е. случилось событие on в сегменте “8”. Если при этом a+b=d, тогда независимо от цвета светофора m_tl необходимо будет ожидать того момента, когда выполнится условие a+b<d. Это условие вставляется в качестве ограничителя в сегменте “9” перед проверкой цвета светофора m_tl.

```
process M_out {
    8: on → on:= false; #9
    9: a+b<d → #10
    10: if (m_tl=green) #12 else #11
    11: c=0 → isl_tl:=red; m_tl:=green; #12
    12: off → off:= false; a:=a+1; #8
}
```

Итак, программы четырех контроллеров построены. В сумме они определяют полную программу управления движением на мосту. Здесь

не потребовался центральный процесс управления светофорами, используемый в предыдущих решениях [1,2]. Однако контроллеры конструировались независимо. Не рассматривалось взаимодействие между ними; не проверялась возможность конфликтов. А конфликты имеются, например, в случае, если после сегмента "6" следующим будет исполняться сегмент "11".

Анализ взаимодействия процессов

Четвертый метод: при разработке системы, составленной из нескольких параллельных процессов со сложным взаимодействием, строится адекватная полная модель поведения системы и реализуется ее анализ прослеживанием исполнения различных комбинаций сегментов от разных процессов с идентификацией конфликтов между процессами. Данный метод — содержательный аналог метода символьного исполнения программы при формальной верификации.

Построение модели начинается с анализа простых сегментов. Это сегменты "2", "4", "7" и "12" в четырех процессах, реализующих модификацию переменных a, b и c. Между этими сегментами нет конфликтов с учетом того, что они являются атомарными. Следовательно, их можно исключить из дальнейшего анализа взаимодействия процессов.

Случай зеленого светофора (isl_tl=green) в процессе Isl_out также простой, не вызывающий конфликтов. Аналогичным является случай зеленого светофора (m_tl=green) в процессе M_out при выполнении условия a+b<d.

Далее рассмотрим ситуации, связанные с условием a + b <= d. Допустим, нарушается условие a+b<d в процессе M_out. В этом случае процесс зависает на сегменте "9" ожидая, когда сегменты других процессов изменят значения переменных, обеспечив истинность данного условия. А автомобиль ждать не будет. При зеленом светофоре автомобиль, находящийся на сенсоре, заедет на мост. Автомобилист не проверяет условия a+b<d, он ориентируется только на показания светофора в соответствии со схемой на рис. 2. В результате нарушается ограничение a + b <= d, что ошибочно.

Предотвратить появление лишних автомобилей на мосту можно лишь установкой светофора m_tl красным в случае a+b=d сразу после пересчета a:=a+1 в сегменте "12". А сегмент "9" становится бесполезным и удаляется. Сегмент "12" модифицируется следующим образом:

```
12: off ⊓ off:= false; a:=a+1;
            {if (a+b=d) m_tl:=red}; #8
```

После данной модификации программы необходимо будет также исследовать поведение системы для условия a+b=d. Допустим, в процессе M_out реализуется условие a+b=d. При этом должно быть m_tl=red, установленное ранее в сегменте "12". При отсутствии сегмента "9" исполнение переходит к сегменту "11". В случае c=0 далее переходим в сегмент "12", что ошибочно. Следовательно, удаленное условие a+b<d должно присутствовать в сегменте "11", который модифицируется следующим образом:

```
11: c=0, a+b<d →
    on:= false; isl_tl:=red; m_tl:=green; #12
```

Рассмотрим наиболее сложную ситуацию, когда в процессе Isl_out при isl_tl=red исполняется сегмент "6" и мост стал пустым (a=0). Если после сегмента "6" следующим будет исполняться сегмент "12" или "11", то на мосту с двух его концов появятся автомобили, что ошибочно. Данный конфликт между параллельно исполняемыми процессами Isl_out и M_out необходимо как-то разрешить. В исходном решении [1] в случае, когда мост становился пустым, возникает угроза бесконечного мгновенного переключения светофоров. Чтобы это избежать, на пустом мосту переключаются светофоры, и происходит ожидание (возможно очень долгое) автомобиля с противоположной стороны. В нашем решении [2] используется сложный процесс LightControl для управления светофорами. Здесь предлагается другой способ: в процессе Isl_out втянуть сегмент "7" внутрь сегментов "5" и "6"; а в процессе M_out втянуть сегмент "12" внутрь сегментов "10" и "11".

После проведения всех модификаций получаем следующую программу:

```
process M_in {
    2: off → off:= false; c:=c-1; #2
}
process Isl_in {
    4: off → off:= false; a:=a-1; b:=b+1; #4
}

process Isl_out {
    5: on → on:= false;
        if (isl_tl=green) {b:=b-1, c:=c+1;#5} else #6
    6: a=0 → isl_tl:=green; m_tl:=red;
            b:=b-1, c:=c+1; #5
}
```

```
process M_out {
   8: on → on:= false; #10
   10: if (m_tl=green)
       {a:=a+1; if (a+b=d){m_tl:=red};#8} else #11
   11: c=0, a+b<d → isl_tl:=red; a:=a+1;
                    if (a+b<d) {m_tl:=green}; #8
}
```

Для данной программы необходимо повторить ее анализ. Обнаруживается следующая особенность. Внезапное включение светофора красным (m_tl:=red) в сегменте “10” вынуждает автомобиль экстренно остановиться. Поэтому оператор m_tl:=red было бы предпочтительней выполнить после регистрации события off в удаленном сегменте “12”. Однако это лишь частично улучшает ситуацию. Проблема остается: как не напугать автомобилиста и при этом надежно гарантировать непопадание на мост следующего автомобиля. Здесь можно было бы использовать желтый цвет светофора.

Верификация в Event-B

Программа управления движением на мосту была переписана на язык Event-B [1]. Программа верифицирована на разных инструментах системы Event-B [6]. Доказаны основные свойства программы методом дедуктивной верификации. Проверено отсутствие дедлоков в двух разных инструментах. В имитаторе детально проверены все варианты исполнения программы. Обнаружена ошибка, исправленная сегментом “13”:

```
13: off → off:= false; #8
```

Ошибка состоит в том, что сигнал off необходимо гасить в случае, когда возврат управления на сегмент “8” реализуется не через сегмент “12”. Обнаружено также рассогласование между работой сенсора и контроллера: при ожидании переключения светофоров выключение сенсора (событие off) должно происходить лишь после появления зеленого сигнала.

Заключение

Опыт разработки программы управления движением на мосту демонстрирует сложность данной задачи. Ошибки, связанные с отсутствием оператора m_tl:=red в случае a+b=d и ошибочным сегментом “9” — не демонстрационные. Они действительно были допущены авторами на стадии автономного построения контроллеров; обнаружены и исправлены лишь на следующей стадии. Ошибки, найденные при верификации инструментами Event-B [6], нетривиальны. Это еще раз доказывает необходимость привлечения инженеров-верификаторов для разработки критических систем управления с высокой ценой ошибки, о чем заявляют многие исследователи. Отметим, что в Новосибирском государственном университете на механико-математическом факультете читается курс «Формальные методы в программной инженерии». Студенты, прошедшие этот курс, далее могут быть обучены на специальность инженера-верификатора.

В данной статье приведены только некоторые методы, применяемые при разработке систем управления. Возможно применение других методов. Это объектно- и аспектно-ориентированное программирование, оптимизирующие трансформации автоматной программы, преобразование в набор сопрограмм и другие.

Доступна расширенная версия данной статьи.

Ссылки

1. Abrial J.-R. Modeling in Event-B: System and Software Engineering. Cambridge University Press, 2010. 586p.
2. Шелехов В. И. Автоматное программирование на базе системы моделирования и верификации Event-B / Программная инженерия. Т. 13, № 4, 2022 — С. 155–167.
3. Тумуров Э. Г., Шелехов В. И. Технология автоматного программирования на примере программы управления лифтом // Программная инженерия. Т. 8, № 3. 2017. С. 99–111.
4. Шелехов В. И. Разработка автоматных программ на базе определения требований // Системная Информатика. Новосибирск: ИСИ СО РАН, 2015. № 1. С. 1–29.
5. Systems and software engineering — Life cycle processes — Requirements engineering. ISO/IEC/ IEEE29148, 2011, 95p.
6. Rodin User's Handbook. Version 2.8 // Jastram M. (editor). 2014. 184p.

МЕТОДИКА ОЦЕНКИ СОСТОЯНИЯ И ВЫБОРА СТРУКТУРЫ ВЫСОКОНАДЕЖНОЙ РАСПРЕДЕЛИТЕЛЬНОЙ СЕТИ

НА ОСНОВЕ ПРИМЕНЕНИЯ МЕРЫ НЕОПРЕДЕЛЕННОСТИ ИНФОРМАЦИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования «Хакасский государственный университет им. Н.Ф. Катанова», Инженерно-технологический институт
655017, Республика Хакасия, г. Абакан, пр. Ленина, 92/1

Дулесов Александр Сергеевич
доктор технических наук, доцент, профессор кафедры цифровых технологий и дизайна

Карандеев Денис Юрьевич
кандидат технических наук, доцент кафедры цифровых технологий и дизайна

Она позволяет повысить эффективность анализа надежности распределительных сетей за счет использования методики вычисления информационной энтропии. Программный продукт предназначен для построения высоконадежных оптимальных структур инженерно-технических сетей с применением меры неопределенности информации, взятой из теории информации.

Карандеев Денис Юрьевич, кандидат технических наук, доцент кафедры цифровых технологий и дизайна

Задача проектирования структур инженерных сетей является на сегодняшний день актуальной, требующей дополнительных научно-практических обоснований. Это обусловлено тем, что в ходе эксплуатации построенных по заранее спроектированным планам структур впоследствии возникают неучтенные издержки, вызванные простоями оборудования сети, обладающей высоким износом.

Теоретическая значимость разработки заключается в исследовании и развитии общего под-

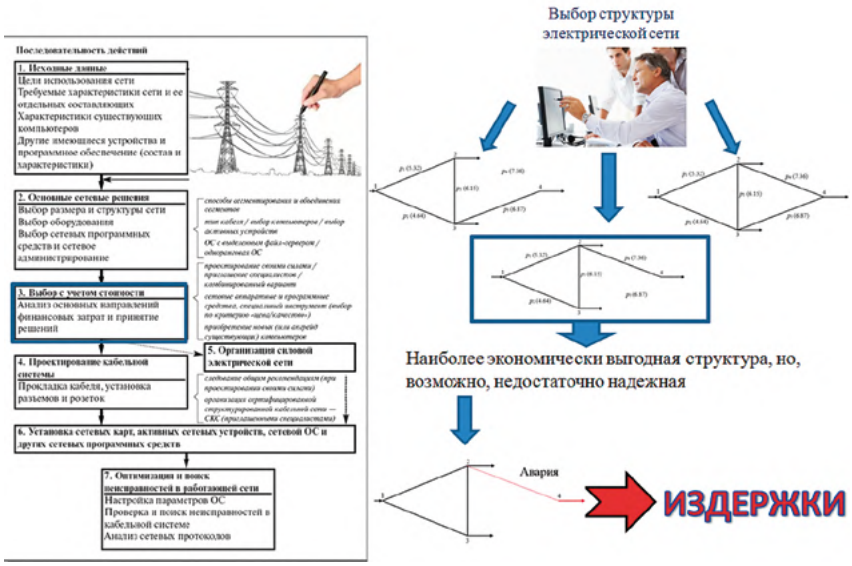


Рис. 1. Этапы проектирования распределительной электрической сети.

хода к оценке состояния и выбора структур высоконадежных распределительных сетей на основе созданной и апробированной методики, включающей в себя расчеты меры неопределенности информации.

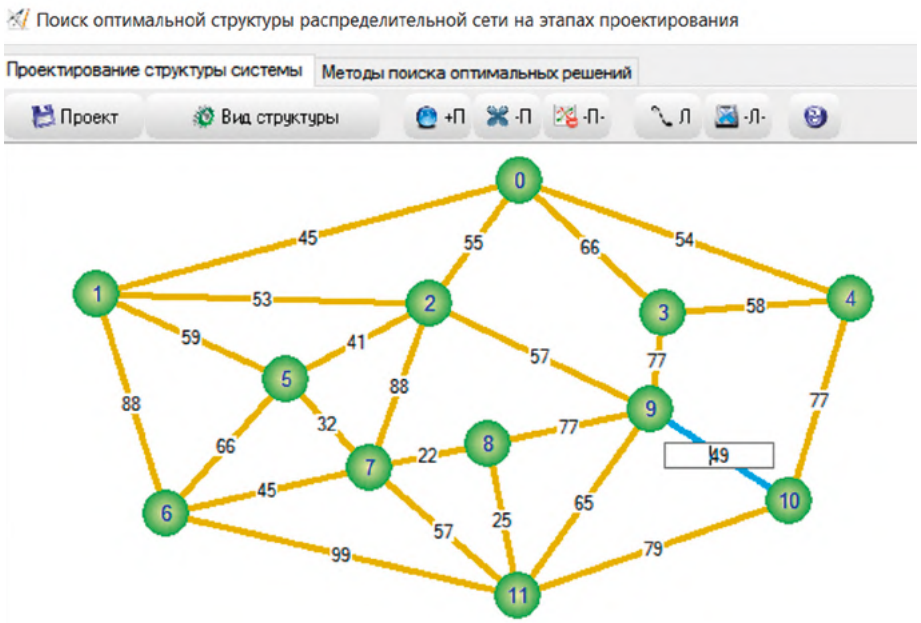
Практическая значимость заключается в применении разработанной методики, когда в исходных данных показателей надежности отражена вероятностная природа поведения распределительных сетей. Тем самым можно рассчитать энтропию, величина которой является характеристикой состояния группы элементов и результатом выбора структуры высоконадежной распределительной сети. Эффекты обусловлены повышением качества анализа надежности сетей, уровня выработки и принятия решений в процессе их проектировании и эксплуатации, что, в свою очередь, снизит издержки в случае перерывов в энергоснабжении.

Авторами предложена научно обоснованная методика описания состояния структуры распределительной сети, относящейся к классу инженерных сетей. В отличие от аналогичной методики, описывающей состояние через учет показателей безотказности, предлагается решение задачи посредством количественного и качественного анализа надежности через численную характеристику информации.

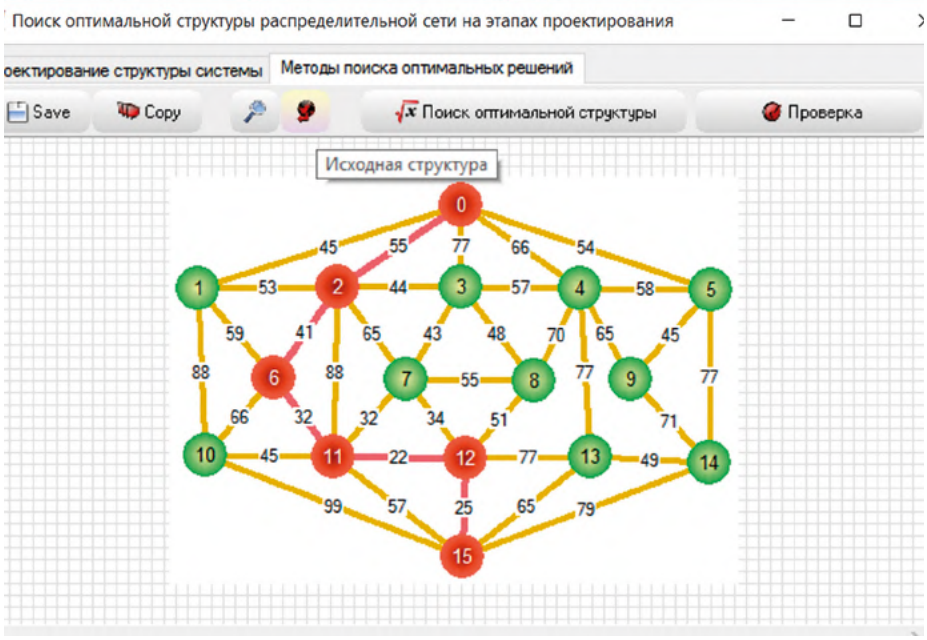
Эффективность предлагаемой методики достигается путем выполнения предлагаемых вычислительных процедур:

- оценка состояний сети и поиска «слабого звена», которая, в отличие от методов анализа характеристик основных событий, позволяет использовать меру неопределенности информации для выявления неблагоприятных событий всех связей структуры;
- расчет оптимального количества однотипных резервных элементов, который, помимо имеющихся методов поиска эффективной структуры, соответствующей требованиям надежности, позволяет учесть технико-экономическую целесообразность функционирования сети;
- выбор структуры замкнутой распределительной сети, отличительной особенностью которого, в сравнении с методами проектирования архитектур систем, является возможность дополнительного учета неопределенности информации, относящегося к статистическим данным и соблюдению условия обеспечения заданного уровня надежности.

Уровень готовности технологии: TRL 7 — прототип системы продемонстрирован в реальных условиях эксплуатации. Свидетельство о государственной регистрации: свидетельство РФ о госрегистрации программы для ЭВМ 2020610366 от 13.01.2020, наименование: «Программа для поиска оптимальной структуры распределительной сети на этапах проектирования».



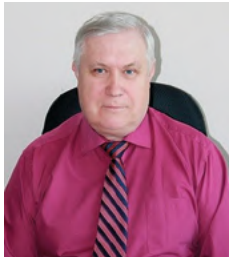
❷



❸

Необходимыми доработками данной методики является включение в методику критерия, учитывающего процесс старения оборудования, что позволит выявить синергетические эффекты.

Одним из подходов к построению методики следует считать применение такой области знаний, как теория информации. Решением задач, касающихся развития теории информации, занимались зарубежные и отечественные ученые: Р. Хартли, К. Шеннон, А. Я. Хинчин, У. Р. Эшби, А. Н. Колмогоров, Д. С. Чернавский, А. Н. Панченков, А. Д. Урсул, М. А. Басин, И. В. Прангишвили, А. М. Хазен, А. В. Мурыгин, Р. Л. Стратонович, В. Гагин, B. Widrow, K. Parthasarathy, B. Hogg, D. Nguyen, Q. Wu, G. Irwin и другие. Вопросами развития и применимости теории надежности распределительных сетей занимались: Б. И. Кудрин, Б. И. Макоклюев, А. Н. Митрофанов, В. З. Манусов, Е. В. Сугак, В. И. Гнатюк, И. И. Надтока, J. Taylor и др. Работы упомянутых ученых позволили обобщить, систематизировать и ло-



Дулесов Александр Сергеевич

Личный комментарий от разработчика:

«Несмотря на имеющиеся эффекты от применения предлагаемой методики, есть обоснование расширить её возможности за счет синергетических эффектов, среди которых — моделирование бифуркационных состояний и процессов, волновых эффектов и квантовых переходов. Это обогатит методику в вопросах формирования и самоорганизации моделей и структур открытых систем, далеких от термодинамического равновесия».

гически увязать исследование о возможности использования меры неопределенности информации, количественная характеристика которой послужила универсальным инструментом решения задач оптимального построения структуры сети.

Аналогами разработанного программного продукта являются программы «Арбитр», разработанные АО «Специализированная инжиниринговая компания Севзапмонтажавтоматика» (СПИК СЗМА) и PSS®SINCAL. Преимущество нашего проектного решения заключается в значительно более низкой стоимости: стоимость ПК «Арбитр» на 1 рабочее место для коммерческих организаций согласно прайс-листу на официальном сайте компании составляет 715 тыс. рублей, для образовательных организаций — 200 тыс. рублей. Стоимость нашего программного комплекса значительно ниже (20 тыс. рублей). Недостатком программы PSS®SINCAL является то, что её разработкой занимается компания SIEMENS (в период активной стратегии импортозамещения целесообразным является использование отечественных разработок) и её деятельность специализируется лишь в проектировании, моделировании и анализе электрических сетей. Наш продукт нацелен на решение задачи поиска оптимальных структур инженерно-технических сетей различного рода. Также к достоинствам нашего продукта можно отнести простоту обучения (интуитивный интерфейс) и инновационную методику анализа надежности (на базе теории информации), которая существенно упростит расчет уровня надежности структуры сети.

На данный момент выявлена заинтересованность в приобретении программного продукта от филиала ПАО «Россети Сибирь» — Хакасэнерго, а также от МУП «Абаканские электрические сети». Ведутся переговоры с энергетическими компаниями в Сибирском федеральном округе.

РОССИЙСКАЯ ФЕДЕРАЦИЯ

СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2020610366

Программа для поиска оптимальной структуры распределительной сети на этапах проектирования

Правообладатель: *федеральное государственное бюджетное образовательное учреждение высшего образования «Хакасский государственный университет им. Н.Ф. Катанова» (RU)*

Автор: *Кирандеев Денис Юрьевич (RU)*

Заявка № 2019667014
Дата поступления 19 декабря 2019 г.
Дата государственной регистрации в Реестре программ для ЭВМ *13 января 2020 г.*

Руководитель Федеральной службы по интеллектуальной собственности

Г.П. Ивлиев

4

Запрос на индустриальное партнерство с целью доработки разработанной методики и ее программной реализации, необходимы инвестиции от потенциального индустриального партнера.

Рис. 2. Проектирование структуры системы.

Рис. 3. Методы поиска оптимальных решений.

Рис. 4. Свидетельство о государственной регистрации программы для ЭВМ.

FIS PLATFORM: LOW-CODE ПЛАТФОРМА ДЛЯ АВТОМАТИЗАЦИИ ФИНАНСОВОГО СЕКТОРА

FIS (Финансовые информационные системы)
630090, г. Новосибирск,
ул. Николаева, 12 (8, 10 этаж)

Кашменский Евгений Никитович
директор и основатель FIS

Ситников Алексей Владимирович
ведущий разработчик FIS

Гончаров Егор Игоревич
ведущий разработчик FIS

Гонтарева Александра Антоновна
PR-менеджер FIS

FIS Platform — ИТ-решение для быстрой цифровизации бизнес-процессов организаций с использованием low-code технологий. С его помощью можно легко выстроить идеальный сценарий пути клиента: от лидогенерации до дальнейшего сопровождения и удержания через все возможные каналы коммуникации.

FIS Platform представляет собой библиотеку готовых модулей и приложений, разработанных на low-code платформе Case Platform, созданной технологическим партнером FIS — компанией Case Studio.

Решение FIS Platform включает в себя:

- среду разработки и отладки;
- библиотеки внутренних и внешних компонентов;
- готовые приложения для автоматизации бизнеса.

С помощью включенного инструментария FIS Platform решает проблему длительной и дорогостоящей цифровизации бизнес-процессов

Управляющий партнер FIS Владимир Залеский на бизнес-завтраке FIS Collection.

Рис. 1. Ключевые клиенты FIS — лидеры финансового рынка.

для финансовых организаций, позволяет стратегически наращивать LTV организации заказчика.

В основе продукта — low-code инструментарий, который позволяет развивать и сопровождать ПО не высококвалифицированными разработчиками, а бизнес-пользователями. То есть аналитики на стороне заказчика могут самостоятельно пройти все этапы создания приложения, включая разработку модели данных, настройку бизнес-процессов, UI, интеграций, печатных форм и отчетности.

FIS Platform предлагает использовать заготовленные шаблоны для быстрой разработки новых решений из смежных отраслей или создать новую систему автоматизации бизнес-процессов любой отрасли с нуля. Компания FIS является победителем премий Retail Finance и European Risk Management Awards в области финансовых технологий.

Платформа разработана компанией FIS, резидентом «Сколково» — опытным отечественным вендором low-code систем, почти 20 лет поставляющим свои ИТ-решения заказчикам на финансовом рынке. 500+ реализованных

Рис. 2. Готовые модели и функции в платформе FIS Platform.

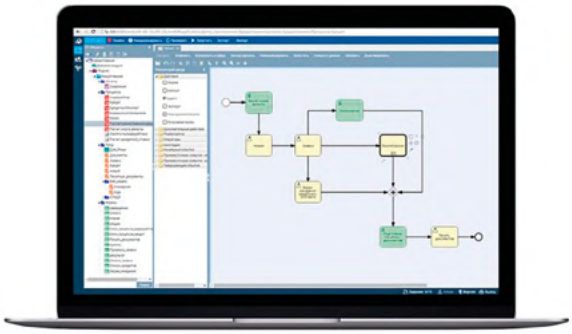


Рис. 3. Дизайнер процессов.

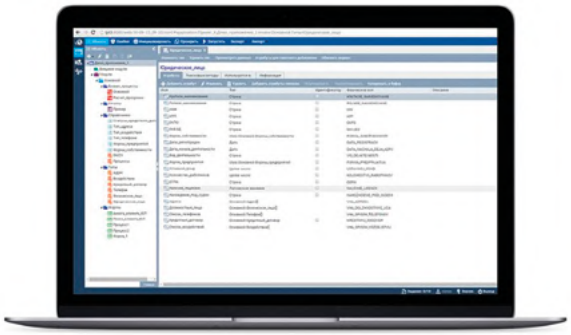


Рис. 4. Конструктор модели данных.

проектов для различных ведущих банков РФ — «Газпромбанка», «Промсвязьбанка», «Совкомбанка», «Открытия», а также многих других. Мы обеспечиваем банкам и другим финансовым организациям надежную непрерывную работу сервисов для полного цикла обслуживания клиентов.

FIS Platform — low-code платформа для оперативной разработки и гибкой настройки отказоустойчивых высоконагруженных корпоративных информационных систем.

Как устроена low-code платформа FIS Platform

FIS Platform — low-code платформа, позволяющая упростить создание и настройку профильных корпоративных приложений с учетом специфики бизнеса. Она включает пакет готовых инструментов для разработки мобильных и веб-приложений, а также удобный визуальный интерфейс, упрощающий их настройку и отладку. С помощью платформы можно значительно сократить время реализации цифровой трансформации бизнеса, упрощая автоматизацию задач заказчика.

Базовая функциональность FIS Platform включает:

- **Дизайнер процессов**

  Визуальный редактор дает возможность упростить моделирование и настройку бизнес-процессов. Он позволяет автоматизировать рутинные, чаще всего повторяющиеся задачи компании, обрабатывать большое количество операций, обогащать их данными из внутренних и внешних источников, управлять интеграционными потоками.

- **Конструктор форм**

  С помощью конструктора форм можно настраивать ПО, не используя программный код. На платформе всегда доступны к использованию уже готовые функциональные блоки, из которых, как из элементов Lego, можно собирать любые интерфейсы и запускать динамические процессы. Например, добавляя в систему всплывающие окна или оповещения, появляющиеся в тех или иных добавленных условиях.

- **Конструктор модели данных**

  FIS Platform также позволяет разрабатывать кастомизированные модели для хранения и обработки данных в системе. К примеру, с помощью встроенного конструктора всего за один клик можно создавать связанные типы данных.

- **Интеграции с различными системами**

  Функциональность платформы можно постоянно расширять. Принцип low-code интеграции позволяет добавлять как сторонние продукты, например внешние базы данных, так и дополнительные сервисы, уже входящие в библиотеку готовых конфигураций FIS.

Результаты внедрений FIS Platform

Платформа FIS Platform — полностью готовое решение, демонстрирующее свою эффективность на протяжении многих лет. С ее помощью уже были созданы сложные высоконагруженные системы для банков из топ-10 РФ. В числе реализованных проектов:

- Уникальная система андеррайтинга и верификации для банка «Открытие». С ее помощью банку удалось в три раза сократить время рассмотрения кредитных заявок, автоматизировать более 40 процессов и увеличить объем обрабатываемых запросов до 100 млн за первые 3 года развития системы, а также подключить к ежедневной работе около 100 пользователей и 40 внешних баз.

- Автоматизированная система для оперативного оформления карт «Халва» и других кредитных продуктов «Совкомбанка». Благодаря платформе в 2020 году была обеспечена онлайн-выдача более 6 миллионов карт «Халва».

- Автоматизация факторинга для «Райффайзенбанка». Уникальный факторинговый проект позволил банку всего за два года войти в топ-7 факторинговых компаний с портфелем финансирования в 17,7 млрд рублей и увеличить число клиентов факторинга в 4 раза.

- Автоматизированная система проверки банка на наличие мошеннических действий для «Почта Банка». Создали систему проверки всех точек продаж банка по набору преднастроенных правил.

- Реализация конвейера автокредитования и системы управления партнерами для банка «Открытие». Обеспечена обработка более 50 тыс. заявок ежедневно, с нуля создана служба верификации, на 20% уменьшено время оформления заявки и на 27% сокращено время выдачи кредитов.

- Полностью автоматизированная система взыскания кредитной задолженности для работы на всех этапах (Pre, Soft, Hard и Legal) для «Газпромбанка». Система FIS Collection обеспечила бесперебойную работу более 700 пользователей при объеме базы данных более 2 млн договоров.

Преимущества FIS Platform

Ключевое преимущество FIS Platform для разработки корпоративных систем — простота. В числе ее основных достоинств:

1. **Скорость разработки и внедрения решений**

   Платформа — это база, на которой из множества готовых функциональных блоков выстраивается цельная корпоративная система. С этими блоками, заранее созданными и прошедшими все тесты, приложению не грозят проблемы с отказоустойчивостью или вертикальной и горизонтальной масштабируемостью.

2. **Быстрая адаптация под новые задачи**

   Low-code позволяет наладить прозрачную связь между бизнесом и IT. Он выступает своего рода метафорой, наглядно демонстрирующей ту или иную модель данных, интерфейс или логику, необходимые бизнесу в определенный момент.

3. **Возможность сэкономить на системе в целом**

   За счет возможности разработки с использованием контроля версий, реализации приложений в микросервисном стиле, а также построения полноценного CI/CD процесса технологическая база FIS Platform может упростить поддержку системы на фоне роста ее сложности.

Актуальность FIS Platform для бизнеса

Сегодня бизнесу важно создавать современные и полностью готовые к использованию продукты в максимально короткие сроки, ведь это позволяет ему оставаться конкурентоспособным. С FIS Platform бизнес получает возможность:

- Быстро разрабатывать, внедрять и выпускать на рынок готовые цифровые решения, которые позволяют отстроиться от конкурентов и выйти на лидерские позиции.

- Оперативно замещать решения зарубежных вендоров, покинувших рынок в 2022 году: Oracle Siebel, SAS RTDM, Microsoft Dynamics и др.

- Получать преимущество за счет кастомизации сервисов, предлагая рынку уникальные бизнес-процессы.

История развития low-code

Первые low-code решения создавались еще в 90-х годах прошлого столетия. Тогда они не стали популярными на рынке, потому что требовали слишком высокой и, как следствие, до-

рогой производительности от систем, в которые внедрялись.

Сегодня программные продукты достигли того уровня развития, на котором не требуют от систем высоких показателей производительности. Поэтому сейчас low-code технологии — как на западных рынках, так и в России — активно развиваются и функционально усложняются. Всего через несколько лет, которые станут прорывными для low-code, рынок разработки будет выглядеть совсем не так, как сегодня:

- По прогнозам Gartner, к 2025 году 70% новых приложений, разрабатываемых предприятиями, будут использовать технологии с low-code.
- Согласно данным аналитического агентства IDC, к 2026 году более 40% компаний по всему миру будут использовать low-code, при этом сокращая усилия по развертыванию сервисов на 33%.
- В исследованиях Expert Market Research рынок low-code разработки вырастет на 26,5% с 2022 по 2027 год.

Иными словами, low-code технологии развиваются и будут развиваться в ускоренном темпе в ближайшие годы по всему миру.

FIS Platform: для автоматизации каких систем подходит

Для автоматизации на базе платформы FIS Platform подходят:

- системы, требующие глубокой кастомизации;
- корпоративные системы для работы внутри компании и коммуникации с партнерами.

Решение подходит для закрытия задач, которые касаются производственно-продуктового цикла компании: документооборота, CRM-систем, систем по работе с закупками/поставками, автоматизации задач маркетинга, систем управления рисками внутри организации и других.

Иными словами, платформа будет полезна для тех компаний, которые, не имея собственного ИТ-подразделения, не сумели подобрать коробочного решения для своих задач и решили закрыть их через внедрение готового инструмента для быстрой разработки. С помощью платформы FIS Platform бизнес в краткие сроки может запустить проект на основе готовых наработок и быстро получить первые наглядные результаты.

Компания FIS заинтересована в партнерах с бизнес-компетенциями в различных отраслях. Такое сотрудничество позволит создать высококонкурентный отечественный продукт, основанный на экспертном опыте и преимуществах FIS Platform. В рамках совместного ведения проекта партнер может полностью делегировать разработку FIS, взяв на себя только управление требованиями к продукту, или вести собственную разработку продуктов на платформе, получив доступ к FIS Platform.

Евгений Никитович Кашменский

Основатель и директор FIS

Личный комментарий от разработчика:

«FIS Platform становится связующим звеном для бизнеса, которое позволяет разным подразделениям корпорации общаться на одном языке. С ее помощью можно автоматизировать любые бизнес-процессы, бесшовно интегрируя дополнительные модули и контролируя все решения с помощью единого центра поддержки».

ФОРВАРД

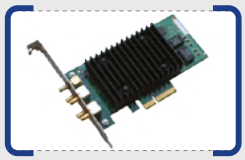
SOFTLAB-NSK

> 30 ЛЕТ НА РЫНКЕ ПРОФЕССИОНАЛЬНОГО ТВ-ОБОРУДОВАНИЯ И ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

NEW

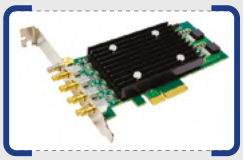
ПЛАГИН ДЛЯ
НОРМАЛИЗАЦИИ ГРОМКОСТИ
В РЕАЛЬНОМ ВРЕМЕНИ
SLSoundLeveler

ПЛАТЫ ВВОДА/ВЫВОДА СЕРИИ FDEXT



FD922
2 входа и 2 выхода;
12G/6G/3G/HD/SD-SDI,ASI

FD722
2 входа и 2 выхода;
3G/HD/SD-SDI, ASI



FD788
до 8 входов/выходов;
3G/HD/SD-SDI, ASI

FD720
2 входа;
HDMI

FD940
4 входа;
HDMI

ПРОДУКТЫ «СОФТЛАБ-НСК» ДЛЯ АВТОМАТИЗАЦИИ ТЕЛЕРАДИОВЕЩАНИЯ

ФОРВАРД Т
автоматизация ТВ-вещания «телеканал-в-коробке»

ФОРВАРД ПЛАГИНЫ
дополнительные опции, расширяющие функционал продуктов

ФОРВАРД ГОЛКИПЕР
замедленные повторы в прямом эфире

КОДЕРЫ/ДЕКОДЕРЫ
продукты для решений с перекодированием ТВ-сигнала

ФОРВАРД СПЛАЙСЕР
бесшовная вставка контента в программы TS

ФОРВАРД РЕФЕРИ
многоканальный сервер системы «Видеогол»

ФОРВАРД ОФИС
управление базой видеоматериалов и программирование эфира

SLADSREMOVER
вырезка рекламы в ретранслируемом сигнале

ФОРВАРД СПОРТИВНЫЕ ТИТРЫ
графическое оформление спортивных трансляций

ТВ-СТУДИЯ ALL¹MIX
интегрированный программный комплекс телевизионной студии

FORWARD4SKYPE
интеграция звонков Skype в передачи прямого эфира

ФОРВАРД ИНЖЕСТ
запись многокамерной съемки

ООО «СофтЛаб-НСК»
+7(383) 363-04-62
sales@softlab.tv
SoftLab-NSK
www.softlab.tv
SoftLabTV
SoftLabTV

ВЕРНУТЬСЯ К СОДЕРЖАНИЮ

Новосибирский центр информационных технологий

unipro.ru

УНИПРО – компания-разработчик ПО, от компиляторов и математических библиотек до биоинформационного софта и бизнес-приложений. Уже 30 лет мы работаем на международном ИТ-рынке, сотрудничаем с высокотехнологичным бизнесом, научными и образовательными организациями.

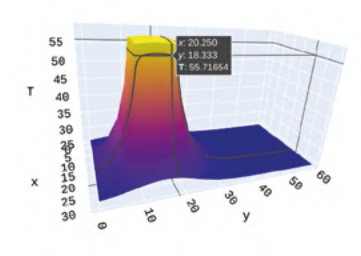
- **КОМПИЛЯТОРЫ И ВИРТУАЛЬНЫЕ МАШИНЫ** – разработка виртуальных машин и компиляторов – основная компетенция компании.
- **МАТЕМАТИЧЕСКОЕ ПО** – библиотеки и системы моделирования.
- **БИОИНФОРМАТИКА** – программные решения для молекулярных биологов.
- **ПОРТИРОВАНИЕ И РАЗРАБОТКА ПОД ПЛАТФОРМУ ЭЛЬБРУС** – создаем и поддерживаем системное ПО, портируем сторонние приложения.
- **ТЕСТИРОВАНИЕ** – тестовые сюиты для языков программирования, анализ производительности, стресс-тестирование и др.

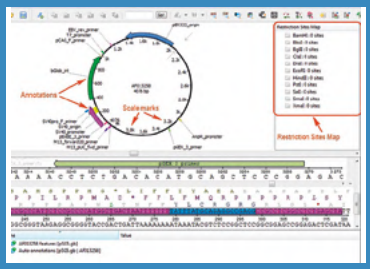
В ПОРТФЕЛЕ СОБСТВЕННЫХ ПРОДУКТОВ:

UGENE (ugene.net) – открытая платформа биоинформатики для работы с генетическими данными.

USPARS – решатель больших систем уравнений для математического моделирования.

UTERM – пакет для моделирования теплообмена в процессорах и других устройствах.





СРЕДИ ЗАКАЗНЫХ НАУКОЕМКИХ ПРОЕКТОВ:

Платформа для моделирования гидроразрыва нефтяного пласта, в сотрудничестве с НГУ.

Программный инструмент для дизайна олигонуклеотидов, сборки и клонирования генных конструкций.

Набор инструментов разработчика для ИИ-специализированного процессора.

marketing@unipro.ru | **тел.: (383) 373-24-63**

ВЕРНУТЬСЯ К СОДЕРЖАНИЮ

МАТЕМАТИЧЕСКИЙ ЦЕНТР В АКАДЕМГОРОДКЕ

ЦЕЛИ МАТЕМАТИЧЕСКОГО ЦЕНТРА В АКАДЕМГОРОДКЕ (МЦА):

- Обеспечение лидирующей позиции Российской Федерации в числе пяти ведущих стран мира, осуществляющих научные исследования и разработки в области математики и смежных областях.
- Обеспечение привлекательности работы в Сибирском макрорегионе для российских и зарубежных ведущих ученых и молодых перспективных исследователей.

НАУЧНЫЕ НАПРАВЛЕНИЯ:

- Пространственные динамические стохастические процессы
- Прикладные цифровые технологии
- Современные математические модели и численные методы ньютоновской механики сплошных сред с применением к геофизике
- Обратные задачи естествознания
- Криптография и информационная безопасность
- Цифровизация математических моделей и интеллектуальные системы обработки данных
- Аксиальные алгебры и связанные с ними группы
- Прикладная абстрактная алгебра: алгебраические методы в топологии и комбинаторике
- Геометрический анализ и его приложения
- Геометрические аспекты математической физики
- Дифференциальные уравнения и динамические системы

Показатели деятельности

200+ Сотрудников	**85+** Ведущих ученых	**>65%** Молодых исследователей
3 Научно-исследовательские программы	**14** Научных проектов	**165** Статей в высокорейтинговых журналах (Q1-Q2) с момента создания

Конференции МЦА

35+ Конференций	**3000+** Участников	**2000+** Молодых участников

Образовательные программы МЦА

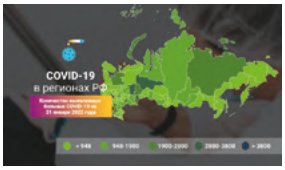
10+ Образовательных программ	**500+** Обученных	**400+** Молодых обученных

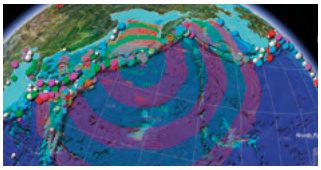
Институт математики им. С.Л. Соболева (научные исследования)

Новосибирский государственный университет (образовательные программы и прикладные исследования)

ВЕРНУТЬСЯ К СОДЕРЖАНИЮ

ОБРАТНЫЕ ЗАДАЧИ ЕСТЕСТВОЗНАНИЯ





ОСНОВНЫЕ РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ:

- Платформа для расчета сценариев развития эпидемий, социальных настроений.
- Программный комплекс моделирования природных и техногенных катастроф, включая накат волны цунами, вулканы, землетрясения.

ОСНОВНЫЕ НАПРАВЛЕНИЯ ИССЛЕДОВАНИЙ:

- Создание и анализ математических моделей взаимосвязанных экономических, эпидемиологических и экологических процессов с учётом социологической ситуации в регионах.
- Машинное обучение, Data-Driven, Physics Informed Neural Networks.
- Прямые и обратные задачи в науках о Земле, жизни и человеке.

МЫ ОРГАНИЗУЕМ КОНФЕРЕНЦИИ

Школа-конференция «Теория и численные методы решения обратных и некорректных задач», Новосибирск

Международная конференция «Современные проблемы обратных задач», Новосибирск

Школа «Обратные задачи и машинное обучение», Сочи, Сириус

BGRS Symposium «Mathematics, bioinformatics and systems computational biology of COVID-19», Novosibirsk

Кабанихин Сергей Игоревич

Научный руководитель группы – член-корреспондент РАН

Криворотько Ольга Игоревна

Руководитель молодежной группы – доктор физико-математических наук

КОМПАНИЯ-РЕЗИДЕНТ «СКОЛКОВО» FIS – ДЛЯ АВТОМАТИЗАЦИИ БАНКОВ И ФИНАНСОВЫХ ОРГАНИЗАЦИЙ

FIS – ОДИН ИЗ ЛИДЕРОВ РОССИЙСКОГО ИТ-РЫНКА

Компания FIS решает проблему длительного и дорогостоящего процесса цифровой трансформации банковского бизнеса. Системы на базе инновационных no-code технологий позволяют:

- оперативно автоматизировать бизнес-процессы без привлечения сторонних IT-специалистов;
- разрабатывать, внедрять, отлаживать ПО и бесшовно интегрировать его с другими продуктами;
- снизить требования к квалификации разработчиков, расширить источники найма и кратно уменьшить затраты на создание систем;
- сократить Time-to-Market, контролировать трансформацию бизнес-процессов с помощью визуальных инструментов и использования Agile.
- создавать новые решения на базе готовых дистрибутивов библиотеки FIS.

500+ реализованных проектов

35+ готовых решений для банков и финансовых организаций

ПРЕИМУЩЕСТВА АВТОМАТИЗАЦИИ БИЗНЕС-ПРОЦЕССОВ С СИСТЕМАМИ FIS

- **Единая платформа для автоматизации любых продуктов**

  Создавайте бизнес-приложения для автоматизации финансового сектора и создания банковских систем быстрее – с меньшим количеством программного года.

- **15+ лет опыта внедрений в отрасли**

  Реализуйте проекты на стабильной и производительной платформе, доказавшей свою эффективность в 100+ крупных проектах.

- **Быстрая разработка с помощью готового набора инструментов**

  В 2 раза ускорьте разработку: используйте единую платформу с проверенным инструментарием для настройки решений под специфику клиента на всех этапах.

ВЕРНУТЬСЯ К СОДЕРЖАНИЮ