

**CPT2018 The International Scientific Conference
of The Moscow Institute of Physics and Technology
of The Institute of Computing for Physics and Technology**

May 28-31, 2018, TzarGrad, Moscow region, Russia

CPT2018 The Conference Proceedings

**СРТ2018 Международная научная конференция
Московского физико-технического института
Института физико-технической информатики**

28-31 мая 2018 г., ЦарьГрад, Московская область, Россия

СРТ2018 Труды Международной научной конференции

©Институт физико-технической информатики, 2018

Москва-Протвино 2018

Комитеты СРТ2018

ISBN 978-5-88835-036-2



Организационный Комитет:

Почетный Председатель – академик РАН Бугаев Александр Степанович,
Президиум РАН, МФТИ, ИРЭ РАН

Андрей Бутко, АО «Русатом АСУ» – Заместитель Председателя Оргкомитета

Георгий Генс, ЛАНИТ – Генеральный спонсор

Мария Берберова, АО «Русатом АСУ» – Ответственный Секретарь

Надежда Шалашилина, ЛАНИТ – Почетный Секретарь

Владимир Захарушкин, АО «Русатом АСУ» – Председатель жюри конкурса студенческих работ

Андрей Клименко, ИФТИ – Технический директор

Владимир Пестриков, МФТИ – Руководитель волонтеров

Рашид Залялов, ИФТИ – Заместитель Председателя Оргкомитета

Равшан Бурхонов, МФТИ – веб-мастер

Программный Комитет

Председатель: Сергей Ротков, ННГАСУ

Члены Программного Комитета:

Алешин Владимир (МФТИ), Астахов Юрий (ИФТИ), Афанасьев Валерий (МФТИ), Бобков Валерий (ИАПУ ДВО РАН), Бондарев Александр (ИПМ РАН), Брусенцев Петр (AVILab), Васин Юрий (ННГУ), Вельбицкий Игорь (Фонд Глушкова, Украина), Венгринович Валерий (ИПФ НАН, Белоруссия), Вольфенгаген Вячеслав (МИФИ, МФТИ), Галактионов В.А. (ИПМ РАН), Дебелов Виктор (ИВМиМГ СО РАН), Дижевский Андрей (ИФТИ), Долговесов Борис (ИАиЭ СО РАН), Ерёмченко Евгений (Неогеография), Жданов Александр (ИТМиВТ), Захарушкин Владимир (МФТИ), Иевлева Ольга (ЮФУ), Исламов Рустам (МФТИ), Кириллов Игорь (МФТИ), Клементьев Александр (США), Клименко Александр (США), Клименко Андрей (ИФТИ), Кружалин Виктор (МГУ, Геофак), Крымский Кирилл (МФТИ), Куницкий Игорь (РДТех), Копылов Андрей (ТулГУ), Лаврентьев Михаил (НГУ), Малофеев Валерий (ПРАО), Массель Людмила (ИСЭМ СО РАН), Местецкий Леонид (МГУ), Минаев Владимир (МВТУ), Муратова Галина (ЮФУ), Никитин Игорь (ИФТИ), Никитина Ляля (ИФТИ), Орлов Дмитрий (ИФТИ, РГ-

ГУ), Орлова Наталья (РИНЦ, eLibrary), Орлов Юрий (ИПМ РАН), Пирогов Юрий (ИФВЭ), Райков Александр (Аналитическое агентство «Новые стратегии»), Рейнгольд Леонид (РДТех), Розман Джон (Белоруссия, США), Романов Алексей (МФТИ), Ротков Сергей (ННГАСУ), Сигов Александр (МИР-ЭА), Слободюк Евгений (МФТИ), Сметанин Юрий (РФФИ), Судейкин Михаил (МФТИ), Столин Вячеслав (ИТЭФ), Тирас Харламий (ПуцГЕНИ), Толок Алексей (ИПУ РАН), Тюльбашев Сергей (ПРАО), Холево Александр (МИ РАН), Цыганов Сергей Алексеевич (РФФИ), Чеповский Андрей (ВШЭ), Юровицкий Владимир (ИФТИ).

IPC Members: Arnold David (UK), Beckhaus Steffi (Germany), Sabine Coquillart (INRIA, France), Gavrilova Marina (University of Calgary, Canada), Wolfgang Heiden (BRS, Germany), Hagen Hans (Germany), Andre Hinkenjann (IVC BRS, Germany), Hui-Huang Hsu (Taipei), Lian Li (HeFei University of Technology, China), Iglesias Andres (University of Cantabria, Spain), Michalcak Georg (Switzerland), Magnenat-Thalmann Nadia (Switzerland), Mura Gianluca (Politecnico Di Milano, Italy), Nam Ling (Santa Clara University, USA), Nielson Gregory (USA), Dimitri Plemenos (France), Paltashev Timour (USA), Prasolova-Forland Ekaterina (NTNU, Norway), Qing Li (City University of Hone Kong), Samet Refik (Ankara Uni, Turkey), Savchenko Vladimir (Hosei University, Japan), Seaman Bill (Duke Uni, USA), Sourin Alexei (NTU, Singapore), Sourina Olga (FhG IGD, Germany), Jiao Ying Shi (P.R.China), Shih Timothy K. (Taipei), Skala Vaclav (Czech Republic), Solomonick Abraham (Israel), Thalmann Daniel (Switzerland), Wetzel Iryna (Intergraph, Switzerland), Wolodchenko Alexander (Germany), Yunjun Gao (Zhejiang University, Hangzhou, P.R.China), Yanchun Zhang (Victoria University, Australia; Fudan University, Shanghai, P.R.China).

Генеральные спонсоры:

Российский фонд фундаментальных исследований

АО «Русатом АСУ»

Корпорация ЛАНИТ

ЗАО РДТех(информационные технологии)

Организаторы и спонсоры:

Институт физико-технической информатики

Московский физико-технический институт

Международный центр по ядерной безопасности

ЗАО РТСофт (средства и системы автоматизации)

Полиграфическое объединение «ПРИНТ-Ателье»

Институт прикладной математики им.М.В.Келдыша РАН
НИИ Научно-технических компьютерных технологий Университета ИТМО
Кафедра инженерной геометрии, компьютерной графики и
автоматизированного проектирования ННГАСУ

Адрес оргкомитета: 142281, Московская область, г.Протвино, Заводской проезд, д.6

Секретариат СРТ2018 Тел: +7(916)5075799; E-mail: rashid.zalyalov@gmail.com

Рекомендовано к изданию Объединённым научно-техническим советом Кафедры и Института физико-технической информатики.

©Институт физико-технической информатики, 2018

Сборник посвящается памяти Станислава Владимировича Клименко



Станислав Владимирович Клименко родился 7 ноября 1941 года в Тбилиси, вырос в Белой Церкви на Украине. В 1966 году окончил Московский физико-технический институт по специальности «Экспериментальная ядерная физика», в 1974 году защитил кандидатскую диссертацию, в 1986 году - докторскую, в 1977 году начал преподавательскую деятельность, которую успешно продолжал всю свою жизнь. Сфера его научных интересов была весьма обширна: системы визуализации и виртуального окружения, информационные системы, ситуационный анализ и системы безопасности, Паутина и Интернет, автоматизация эксперимента и анализ данных в ядерной физике.

С 1967 по 2002 годы работал в Институте физики высоких энергий (ИФВЭ) в качестве младшего научного сотрудника, затем руководителя группы, старшего научного сотрудника, начальника лаборатории с возложением обязанностей заместителя начальника отдела, главного научного сотрудника. С 2003 по 2006 годы трудился в Институте космических исследований РАН в должности главного специалиста. С 2006 по 2008 годы в ОАО «Научно-исследовательский центр электронной вычислительной техники» на должности начальника отдела. С 2008 по 2009 годы работал в Политехническом музее в должности ведущего научного сотрудника. С 2009 по 2010 годы в Московском радиотехническом институте РАН в должности начальника сектора. С 2010 по 2011 годы в Институте истории естествознания и техники им. С.И.Вавилова РАН в должности главного научного сотрудника. В октябре 2012 года был избран профессором кафедры инженерной геометрии, компьютерной графики и автоматизированного проектирования Нижегородского Государственного Архитектурно-Строительного Университета (ННГАСУ), являлся председателем диссертационного совета Д 999.048.02 при ННГАСУ.

В МФТИ преподавал с 1977 года: более 15 лет был заместителем заведующего базовой Кафедры физики высоких энергий, с 1998 по 2011 годы заведующим Кафедрой системной интеграции и менеджмента, в последнее время был научным руководителем Кафедры физико-технической информатики с базовой организацией АО «РАСУ» - «Русатом - Автоматизированные системы управления».

Под руководством С.В. Клименко и под эгидой ИФВЭ-МФТИ в течение 10 лет действовала Школа для молодых ученых «Автоматизация и эффективность научных исследований». Под руководством С.В. Клименко было защищено более 40 дипломных и магистерских работ, 15 кандидатских диссертаций. Был приглашённым профессором Южно-Чешского университета в Ческе-Будеёвице, Наньянского технологического университета в Сингапуре и Тамкангского университета на Тайване, приглашённым исследователем во Фраунгоферовском Институте медиакоммуникаций, Санкт Августин, Германия.

Принимал участие в проектировании, создании и исследовании систем автоматизации экспериментов в физике высоких энергий. Под его руководством выполнен ряд международных проектов по тематике новых информационных технологий двойного назначения, поддержанных грантами министерства науки, высшей школы и технической политики Российской Федерации, РФФИ и Европейского Союза.

Лауреат премии Совета министров СССР по науке и технике 1984 года, ветеран атомной энергетики и промышленности с 1999 года.

В 2001 года под руководством С.В. Клименко была создана первая в России установка виртуального окружения и первая в мире - на одном персональном компьютере. Установка демонстрировалась на Международной выставке СеВIT'2002 и на юбилейной конференции-выставке, посвященной 10-летию РФФИ. Опубликовал более 300 научных работ и 6 монографий.

Был членом секции «Микроэлектроника, средства вычислительной техники и автоматизации, ядерное приборостроение» Научно-технического Совета №1 Минатома России, председателем рабочей группы по конверсии предприятий Минатома «Информатизация научных и проектных работ», выбранным членом НТС ИФВЭ, членом Большого Учёного Совета МФТИ, членом редколлегии журнала Computer & Graphics (Elsevier).

Являлся членом редколлегий журналов LNCS-Springer, «Программирование», «International Virtual Reality Journal», председателем конференций VEonPC20xx, MEDIAS20xx, СРТ20xx, SCVRT20xx, GraphiCon-2015, IEEE Ubi-Media-Computing-2016. С.В. Клименко являлся членом международных профессиональных организаций: Американского математического общества - AMS, Ассоциации вычислительных машин - ACM - senior member, Института инженеров по электротехнике и электронике - IEEE -senior member.

Являлся экспертом РФФИ по направлениям: «Ориентированные фундаментальные исследования» и «Создание и развитие информационных и телекоммуникационных ресурсов для проведения фундаментальных исследований», а также экспертом Европейского союза по информационным технологиям, принимал участие в оценке научной работы Фраунгоферовского института медиакоммуникаций в составе международной группы экспертов, зарегистрирован в Федеральном реестре экспертов научно-технической сферы при Минобрнауки России. Являлся членом Экспертного совета РФФИ по информационным системам.

Станислав Владимирович был основателем Института физико-технической информатики. ИФТИ создавался как независимый научно-исследовательский институт, деятельность которого направлена на развитие фундаментальных междисциплинарных исследований на основе современных компьютерных технологий и их внедрение в России для решения широкого класса задач науки, промышленности и образования. Институт был организован в 1994 году на базе лаборатории системного анализа ИФВЭ в подмосковном Протвино как Российский центр физико-технической информатики (РЦФТИ) по инициативе и при участии Министерства РФ по атомной энергии, Министерства науки и технической политики РФ, МГУ имени Ломоносова, Московского физико-технического института, Московского авиационно-технологического института имени Циолковского, а также ряда предприятий и коммерческих компаний. В 1999 году РЦФТИ был преобразован в автономную некоммерческую организацию «Институт физико-технической информатики».

Институт ведет широкий спектр исследований в области физико-технической информатики, прикладной математики и физики, компьютерной инженерии и системной интеграции. Основные научные направления ИФТИ – системы визуализации и виртуального окружения, инженерная графика и мультимедиа для задач мониторинга и безопасности, стойкость сложных социотехнических систем. В ИФТИ работает более 60 ведущих российских специалистов. С 1995 по 2017 год было выполнено более 100 проектов, поддержанных грантами Российского фонда фундаментальных исследований для решения задач в науке, медицине – например, для проведения сложных операций, промышленности, образовании, музыке, атомной и космической отраслях.

Станислав Владимирович был большим оптимистом и всегда жил с планами о будущем науки, образования и страны. Так одна из его идей - создание элитного научно-исследовательского университета в наукограде Протвино, в идеале – обязательно с кампусом.

Он считал, что основная миссия ученого в обществе – это учить студентов. И в самом деле вся наука развивается исключительно за счет того, что ученые работают вместе со студентами старших курсов и аспирантами. Потому что вся наука делается молодыми людьми, это вообще сфера деятельности молодых, которых опытные ученые должны направлять, координировать. Так устроен человек: наш мозг способен производить что-то новое, творческое, а также воспринимать новые вещи только в молодом возрасте – приблизительно до 35 лет . В научных статьях по нейрофизиологии говорится, что это связано со спецификой головного мозга и его способностью к функциональному развитию.

Станислав Владимирович состоял в браке, отец двоих сыновей и дед четырех внуков. Активно занимался спортом - горными лыжами, плаванием, бегом. В свободное от работы время занимался подготовкой инструкторов горнолыжного туризма во Всесоюзной школе инструкторов, также известной как школа Л.П. Ремизова, следствием чего стало появление современного горнолыжного тренажера в МФТИ.

Все знавшие Станислава Владимировича отмечают его титаническую работоспособность, прирожденные качества лидера, надежность и отзывчивость. Светлая память о нем останется в сердцах его друзей и коллег на всю оставшуюся жизнь.

Дорогие коллеги!

Минул год со дня нашей традиционной международной конференции СРТ2017. Наша конференция, как всегда, стремится мыслить красиво, не разрушая прошлое, опираясь на настоящее, и с явным ориентиром на будущее. Доклады на нашей конференции научно обоснованы, опираются на весомый научно-исторический опыт, экспериментальный базис и незыблемые фундаментальные закономерности.

Импортозамещение требует от страны самых прорывных технологий и амбициозных проектов, над которыми, кстати, активно работают не только ученые России, но и их коллеги из многих стран мира.

Один из представленных на нашей конференции докладов «Разработка методов статистического анализа данных для оценки характеристик «визуального восприятия – действия» в системе виртуального окружения тренажеров» затрагивает тему обработки изображения мозгом и его стабилизации.

Работа «Мультиинструментальный интерфейс для работы с цифровыми голограммами в виртуальном окружении (виртуальной реальности)» нацелена показать актуальность виртуальной реальности, как средства взаимодействия с данными получаемыми методами цифровой голографической микроскопией.

Если мы составляем словарь, то располагаем все помещенные в нем слова по алфавиту соответствующего языка. Это - глобальный принцип словарной работы в любой стране. Он не подходит для собраний образов, которые как и слова пытаются представить и объяснить некие стоящие за образом вещи, состояния либо события. Приходится искать иной способ представления образов в различных их сочетаниях. Этому посвящена статья А.Б.Соломоника «Человеко-знаковые интерфейсы (отчет по итогам годичной работы над созданием «Словаря универсальных по содержанию образов)». В результате множества попыток и на основании прежнего опыта исследований в этом направлении был сделан вывод, что для больших собраний образов организующим началом может быть только скрытая за образом идея, которая будет одинаково понятна в любой культурной традиции любого народа на земле, хотя она всякий раз будет выражена по-своему.

Информационные процессы происходят в мире знаков. В них так или иначе соучаствуют люди. Это значит, что когнитивные модели, которые в них реализованы, так или иначе базируются на семиотических т. е. гуманитарных технологиях или универсалиях. В последние годы, особенно с появлением соцсетей, все большую значимость начинают играть модели и универсалии, основанные на науках о коммуникациях, и прежде всего на древнейшей из них - риторике. Проблемы применимости этих моделей к информационной практике обсуждаются в цикле статей, написанных нашими студентами под руководством к.фил.н., доцента Рыкова В.В.

АСУ ТП – это целый комплекс решений, обеспечивающих автоматизацию основных операций технологического процесса, в частности, на АЭС. Направления развития АСУ ТП, требующие решения задач новыми методами визуализации обрабатываемой информации и обучения персонала, освещены в работах нашего аспиранта Хламова М.А. под руководством к.т.н., доцента Захарушкина В.Ф.

Интересна работа «Технология анализа слабоформализованных данных мультисенсорных систем с применением методов распределенной экспертизы» (авторы д.т.н. Захарова А.А. и к.т.н. Подвесовский А.Г.). Здесь представлена технология адаптируемого многоуровневого анализа слабоформализованных данных мультисенсорных систем, когда источники информации характеризуются гетерогенностью, а принятие решений носит междисциплинарный характер. Рассмотрена структура процесса анализа данных на разных уровнях их представления. Для реализации отдельных этапов анализа данных предложена распределенная экспертиза, предоставляющая возможности территориально распределенного взаимодействия экспертов в разных областях, а также механизмы обеспечения согласованности и сходимости результатов экспертного оценивания.

Не меньший интерес представляет работа наших студентов Берзина В.И. и Какурникова Н.А. «Разработка методов управления движением двухколесного мобильного робота». В работе рассматриваются методы планирования и управления движением мобильного робота. Моделируется исполнительная система на примере двухколесной приводной платформы с массивом курсовых дальномеров.

При автоматическом анализе текстов на наличие предмета интереса существует проблема идентификации его наличия в сложных случаях, особенно когда возможны несколько толкований. Предполагается также, что автоматическая система распознавания справляется с анализом большинства текстов и может выделять относительно редкие случаи, где «интеллекта» системы не достаточно для надежной идентификации искомого смысла в тексте. Эти случаи должны передаваться системой сертифицированным экспертам или их группам для дальнейшего анализа. В этом случае возникает две проблемы. Первая заключается в том, что требуется информация об интеллектуальном потенциале экспертов и, соответственно, о методологии получения этой характеристики. Вторая проблема заключается в справедливой оценке вклада каждого эксперта в общую работу. Решению этих проблем посвящен цикл статей, написанных группой специалистов с участием наших студентов, под руководством к.т.н. Шарнина М.М.

Опасна ли ядерная энергетика? Этим вопросом особенно часто стали задаваться в последнее время, особенно после аварий на известных атомных электростанциях. И если опасность все же имеется, то каким образом можно уменьшить риск неприятных последствий аварии? И где же причина того или иного фактора опасности? По линии упреждения угроз ведутся оригинальные исследования, результаты которых представлены в цикле статей, написанных студентами МФТИ и МГУ им. Ломоносова под руководством к.т.н., доцента Берберовой М.А.

Важные и интересные проблемы подняты в статьях коллектива авторов под руководством д.т.н., профессора Роткова С.И. В частности, обсуждается проблема автоматизированного синтеза предварительной каркасной модели трёхмерного объекта по техническому чертежу, содержащему разрезы и сечения. Описанный процесс позволяет расширить возможности алгоритма синтеза электронной модели 3D объекта по данным технического чертежа, разрабатываемого нижегородской научной геометрической школой под руководством

профессора Роткова С.И. Сложности формализации анализа изображений, составляющих содержание технического чертежа, накладывают ограничения на входную информацию алгоритмов, решающих задачу синтеза, поэтому работа по расширению классов входной информации является актуальной и востребованной.

Нельзя забывать о том, что наиболее весомый вклад в такое благое и полезное дело, как проведение таких конференций и создание таких Сборников, внес научный руководитель кафедры Физико-технической информатики, профессор, Почетный Профессор МФТИ, доктор физико-математических наук Станислав Владимирович Клименко. Его светлой памяти посвящаем мы и нашу Конференцию СРТ2018, и этот Сборник!

Как сказал Поэт: «Безумству храбрых поем мы песню!» В добрый путь, читатель, участник нашего сообщества, или просто неравнодушный к вопросам комплексного освоения прошлого, настоящего, и, самое главное, – будущего.

От имени комитетов:

Академик Бугаев А.С.

Профессор Ротков С.И.

Профессор Райков А.Н.

Содержание

<i>Клименко С. В., Шарнин М. М., Галина И. В., Демидов А. О.</i> Методика построения индекса контекстного научного цитирования (ИКНЦ)	1
<i>Протасов В. И., Потапова З. Е., Шарнин М. М., Колесникова К. А.</i> Определение интеллектуального потенциала и оценка вклада экспертов при использовании системы коллективного согласования решений в сложных случаях анализа текстов	12
<i>Клименко С. В., Шарнин М. М., Галина И. В., Соколов Е. Г.</i> Индекс идеологического влияния: базовые модули комплексной методики	18
<i>Шарнин М. М., Ищенко Н. С., Пахмутова Н. Ю., Сюракшина Ю. В.</i> Использование методов тематического моделирования в мультязыковых средах для прогноза тревожных событий	33
<i>Золотарев О. В., Шарнин М. М., А. Еромасова, Ф. М. Тезадов</i> Современные подходы к обработке многоязычных текстов, основанные на методах дистрибутивной семантики	43
<i>Хакимова А. Х., Шарнин М. М., Кузнецов К. И., Сюракшина Ю. В.</i> Методы оценки семантического сходства как средства обнаружения межъязыкового плагиата	48
<i>Соколов Е. Г., Шарнин М. М.</i> Применение LSA/LSI методов для извлечения данных из неструктурированного текста	59
<i>Кудрявцев Д. В., Алешин В. П.</i> Разработка методов статистического анализа данных для оценки характеристик «визуального восприятия - действия» в системе виртуального окружения тренажеров	64
<i>Киселев В. А., Пигарёв В. А.</i> Мультинструментальный интерфейс для работы с цифровыми голограммами в виртуальном окружении (виртуальной реальности)	73
<i>Тирас Х. П., Киселев В. А., Клименко С. В.</i> Перспективы разработки программно-аппаратного комплекса для оперирования биологических макрообъектов в среде виртуальной реальности	79
<i>Клименко С. В., Пестриков В. И., Чернецкий А. М.</i> Современные методы построения трехмерной модели операционного поля по видеоряду	84
<i>Соломоник А. Б.</i> Человеко-знаковые интерфейсы (отчет по итогам годичной работы над созданием «Словаря универсальных по содержанию образов»)	96
<i>Рыков В. В.</i> Когнитивная информатика и гуманитарные технологии	105

<i>Белова М. С., Рыков В. В.</i> Мультимедийные коммуникации — критерии классификации	110
<i>Рыков В. В. Теренин Р. В.</i> Когнитивные технологии в мультимедийной коммуникации — основные термины и понятия	113
<i>Климчук А. И., Рыков В. В., Сжлярова Е. В.</i> Когнитивные технологии и мультимедийные корпуса	115
<i>Максимов Е. С., Рыков В. В.</i> Когнитивные модели в инновационной деятельности	118
<i>Голубчиков Ю. Н., Клименко С. В.</i> О проблемах имиджевого возвышения России	121
<i>Судейкин М. И., Афанасьев А. А.</i> Разработка и реализация программного обеспечения «Геоконвертор» в составе геоинформационной системы «Геоанализ»	129
<i>Захарушкин В. Ф.</i> Анализ особенностей АСУ ТП АЭС при решении задач демонстрации функциональных возможностей и обучения пользователей	152
<i>Захарушкин В. Ф., Хламов М. А.</i> Выбор аппаратно-технических средств для реализации системы дополненной ре- альности для задач АСУ ТП	164
<i>Завьялов Д. А., Захарова А. А.</i> Уровневое моделирование на основе комплексного подхода	172
<i>Агапов А. Ю.</i> Анализ производительности архитектуры ARM и другой перспективной архитек- туры на SPEC тестах скомпилированных свободной коллекцией компиляторов GNU	176
<i>Злобин Н. Н.</i> Исследование Ni ₃ Nf как легирующего компонента для создания высокотемпера- турного сплава с эффектом памяти формы на основе Ni ₃ Ta	194
<i>Берзин В. И., Какурников Н. А.</i> Разработка методов управления движением двухколесного мобильного робота	207
<i>Башков А. С.</i> Выявление проблем портала государственных услуг на основе анализа текстов обращений пользователей	212
<i>Захарова А. А., Подвесовский А. Г.</i> Технология анализа слабоформализованных данных мультисенсорных систем с применением методов распределенной экспертизы	221
<i>Хаританович Е. А., Оздоева А. Х.</i> Оценка транспортного риска в проектах нефтегазовой промышленности	230

<i>Оздоева А. Х.</i> Оценка эффективности применения технологий, направленных на снижение негативного воздействия производственных процессов нефтегазовой отрасли на окружающую среду	245
<i>Маринина Д. А., Берберова М. А.</i> Исследование зависимости результатов оценок радиационного риска АЭС с реактором типа ВВЭР от состава населения, проживающего вокруг АЭС (на примере Ростовской и Калининской АЭС)	255
<i>Оздоева Л. Х.</i> Геотектоника, как причина различия результатов сравнительной оценки внутреннего и внешнего облучения и ущерба населению от АЭС	264
<i>Киданова А. Ю., Берберова М. А.</i> Исследование влияния воздействия радиоактивных веществ на результаты оценки экологического ущерба вокруг АЭС на примере Калининской АЭС	271
<i>Сынгаевская А. А., Берберова М. А., Дервянжин А. А.</i> Исследование влияния стресса на вероятность совершения ошибок персоналом АЭС в рамках анализа надежности персонала	282
<i>Чернявский К. И., Быков Н. К., Золотарев С. С., Берберова М. А.</i> Разработка атласа оценок показателей риска для российских АЭС	289
<i>Конопацкий Е. В.</i> Геометрическое моделирование и оптимизация многофакторных процессов и явлений методом многомерной интерполяции	299
<i>Дергунов В. И., Лагунова М. В., Смычѣк М. М., Тюрина В. А.</i> Формирование предварительной каркасной модели 3D-объекта по техническому чертежу, содержащему разрезы и сечения	307
<i>Васин Ю. Г., Васин Д. Ю., Громов В. П., Ротков С. И.</i> Робастная векторизация графических документов с выраженной ориентацией линейных объектов	313
<i>Васин Ю.Г., Васин Д.Ю., Громов В.П., Кустов Е.А., Ротков С.И., Ясаков Ю.В., С.В.Клименко</i> Технология создания графических баз данных поэтажных планов зданий	318
<i>Лагунова М. В., Мошкова Т. В., Тюрина В. А.</i> Геометрическое конструирование на параметрической основе с элементами эвристического поиска	321

Методика построения индекса контекстного научного цитирования (ИКНЦ)*

S. V. Klimenko¹, *M. M. Sharnin*², *I. V. Ghalina*², *A. O. Demidov*³
mc@keywen.com

Протвино, ¹АНО «Институт физико-технической информатики»

Москва, ²Институт Проблем Информатики ФИЦ ИУ РАН

³Московский физико-технический институт (государственный университет)

Рассматривается методика автоматизированного построения нового показателя качества научной статьи — индекса контекстного научного цитирования (ИКНЦ); связь ИКНЦ и индекса научного цитирования (ИНЦ) с другим авторским показателем — мерой подобия (мерой семантического сходства) двух произвольных текстов. Приведены результаты экспериментов с этими показателями, в частности изучена корреляция между ИКНЦ и ИНЦ, зависящая от значения порога семантического подобия. На основе моделирования значений независимых переменных и их коэффициентов регрессии предложена прогностическая математическая вероятностная модель зависимости количества прямых цитирований от количества неявных ссылок и их параметров.

Ключевые слова: индекс контекстного научного цитирования, семантическое моделирование, тематическое моделирование, дистрибутивная семантика, семантический поиск, скрытое размещение Дирихле, Word Mover's Distance, нейронные сети, Word2Vec, мера семантического сходства, математическая статистика, большие данные, автоматизированные системы извлечения знаний.

The method for construction of science contextual citation index (SCCI)*

S. V. Klimenko¹, *M. M. Charnine*², *I. V. Ghalina*², *A. O. Demidov*³

Protvino, ¹ANO "Institute of computing for physics and technology"

Moscow, ²FRC CSC RAS

³Moscow Institute of Physics and Technology (State University)

The study is devoted to the method of automated building a quality indicator for scientific articles — the index of contextual scientific citation (ICSC). We explore the relationship between the ICSC, the scientific citation index (ISC) and another author's measure — a semantic similarity of two arbitrary texts. The results of the experiments with these parameters are provided. In particular, they are focused on the analysis of the correlation between the ICSC and the ISC, which depends on the value of the semantic similarity threshold. On the basis of modeling the independent variables and their regression coefficients, a predictive mathematical probability model is proposed. It reflects the dependence of the direct citations' number on the number of implicit references and their parameters.

Keywords: index of contextual scientific citation, semantic modeling, topic modeling, distributive semantics, semantic search, Latent Dirichlet Allocation, Word Mover's Distance, neural networks, Word2Vec, measure of semantic similarity, mathematical statistics, big data, automated systems of knowledge extraction.

Введение

Настоящая статья представляет часть проекта по исследованию и разработке семантических методов построения индекса контекстного научного цитирования (ИКНЦ). Цель проекта — задание теоретических и технологических основ автоматизации получения нового показателя качества научной статьи — ИКНЦ, способного прогнозировать будущие значения стан-

дартного индекса научного цитирования (ИНЦ), который рассчитывается по количеству прямых библиографических ссылок на статью. Проект включает создание корпусов естественно-языковых текстов, относящихся к выбранным предметным областям, на основе обработки большого потока данных из Интернет для поиска дополнительных неявных связей между документами (неформальных ссылок). В рамках проекта были созданы методика, пошаговые алгоритмы и программные средства для построения автоматизированной системы выявления неявных связей между докумен-

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 16-07-00756, 18-07-20030, 18-37-10008.

тами; разработаны алгоритмы расчета индекса контекстного научного цитирования (ИКНЦ). Проведён ряд экспериментов на выделенных тематических коллекциях.

Методика оценки качества научных статей разрабатывается на основе вероятностной модели влияния (impract) научной статьи на ссылки и идеи в последующих статьях, а также на основе модели представления идеи в виде множества похожих по смыслу фраз [1, 2, 3, 4]. Первая попытка применения ИКНЦ состоялась на корпусе по компьютерной графике (КГВ) [5]. ИКНЦ рассчитывается автоматически по неявным контекстным ссылкам на статью и связан со статистической вероятностью ожидаемого появления прямых библиографических ссылок. ИКНЦ обладает прогнозными свойствами, позволяющими делить новые статьи на группы и ранжировать их по качеству. Неявные ссылки в статье — это упоминания чужих идей и их авторов. ИКНЦ полезен для обнаружения новых документов (научных статей), так, чтобы наиболее значимые документы могли быть представлены в большем количестве. Обычно для обнаружения значимых научных документов используется стандартный индекс научного цитирования (ИНЦ), но через ИНЦ невозможно проанализировать новые документы с нулевым значением ИНЦ. ИКНЦ же способен их анализировать, потому что он более чувствителен, чем ИНЦ, так как учитывает неявные ссылки на статью.

Для выявления дополнительных неявных связей/ссылок между документами в ходе выполнения проекта были построены и постоянно пополняются текстовые корпуса. Данные корпуса полезны также для расчета семантических отношений между терминами. Сбор текстов осуществлялся из открытых источников, в том числе из русской Википедии, из социальных сетей, а также проводился поиск статей по связям цитирования в Google. За счёт круглосуточной работы двух серверов постоянно увеличивается объём корпусов, из которых были выделены подклассы экспериментальных (обу-

чающих) выборок по темам «Компьютерная графика, визуализация и виртуальное окружение в СССР и России (КГВ)», «Автономные необитаемые подводные аппараты» (АНПА), «Политический экстремизм» и др. Полученный корпус, составлявший в 2016 более 100 Гб (в том числе по тематике КГВ более 10 Гб текстов, из которых было выделено ядро, содержащее 350 Мб различных неповторяющихся фраз), в 2017 был доведён до объёма в 240 Гб. Так, по теме АНПА был получен общий словарь словосочетаний объёмом 53000 терминов, из которого выделен словарь ключевых словосочетаний в 870 терминов. Размер коллекции-2016 по тематике КГВ, составлявший почти 1000 статей, из которой был выделен экспериментальный корпус в 120 связанных между собой статей, в 2017 вырос более чем вдвое, а экспериментальный корпус расширился до 250 статей, связанных библиографическими ссылками [6]. Был создан экспериментальный корпус по теме «Политический экстремизм»; с помощью системы IQBuzz построены тематические корпуса, например корпус-1: iqbuzzK (целевой запрос «кафиры»), включающий 100 тысяч документов и 25 млн слов, и корпус-2: iqbuzzV (целевой запрос «ватники»), содержащий около 200 тысяч документов и 70 млн слов [3]. Тексты для них собирались из социальных сетей: vk.com, odnoklassniki.ru, livejournal.com, facebook.com, twitter.com и пр.

Методика автоматизированного построения ИКНЦ

Методика базируется на учёте явных и неявных ссылок. Исследование параметров неявных ссылок опирается на гипотезу о том, что на статьи с новыми идеями можно найти много неявных ссылок, новаторские, «прорывные» статьи имеют повышенную вероятность прямого цитирования и более высокие значения ИНЦ.

Вероятностная математическая модель зависимости количества прямых цитирований от количества неявных ссылок и

из параметров строится на основе лингвистического процессора, выделяющего неявные ссылки; процессор настраивается с помощью метода машинного обучения так, чтобы корреляция между индексами ИНЦ и ИКНЦ была максимальной. Базой для разработки принципов построения неявных ссылок между документами является подход, позволяющий осуществлять поиск релевантной информации не только по ключевым словам, но и по ассоциациям, опосредованным отношениям. В основе предлагаемого подхода лежит представление о том, что идеи могут быть адекватно выражены множеством сходных по смыслу фраз или множеством терминов (близким аналогом предлагаемого подхода выражения идей/тем является один из методов тематического моделирования: метод латентного размещения Дирихле — LDA).

В создаваемой авторами проекта базе данных текстовых коллекций необходимо решить задачи определения семантического подобия документов и фраз, выявления семантических связей между терминами, а также классификации документов по категориям в автоматизированном режиме. Смысловое сходство фраз определяется с помощью грамматических трансформаций, переводных эквивалентов, синонимии и функциональной синонимии, а также с помощью обнаружения ассоциативных связей, выявленных по авторской методике построения ассоциативного портрета предметной области (АППО). Для структуризации текстов предметной области и построения иерархии категорий используется авторская методика построения ассоциативно-иерархического портрета предметной области (АИППО), в которой для расчета иерархических связей между значимыми фрагментами текста используются методы тематического моделирования, такие, как LDA. Для улучшения параметров иерархической кластеризации терминов тематическое моделирование применяется совместно с моделями дистрибутивной семантики и семантических векторных

пространств. Организация коллекции текстовых документов в виде кластеров производится так, чтобы документы в пределах каждого кластера были похожи. Текст обычно отображается в векторном пространстве (в виде «мешка слов»), и каждый документ представляется как вектор признаков с использованием схемы взвешивания. Затем выполняется кластеризация путем измерения расстояния между векторами признаков. Выделенные по указанным методикам ассоциативные и иерархические связи между значимыми словосочетаниями и терминами позволяют разрабатывать более совершенные методы и метрики/меры подобия научных текстов. Создаются тематические коллекции и выборки документов, связанных формальными библиографическими ссылками, а также содержащие сходные тексты, связанные неформально. Неявные связи обнаруживаются с помощью меры семантического подобия (введённый авторами показатель).

Краткое описание методик включает в себя следующие этапы:

1. Семантический поиск по заданным ключевым словам и накопление базы релевантных интернет-текстов, из которых формируются корпуса ЕЯ-текстов различных ПО (big data). Сбор текстов осуществлялся в том числе из русской Википедии, из социальных сетей, а также проводился поиск статей по связям цитирования в Google. Данный пункт подразделяется на следующие подпункты (в зависимости от цели исследования):
 - (а) Выбор некоторого (первоначального) набора ключевых слов вручную и поиск по ним в Интернете с целью создания тематической обучающей выборки для последующего машинного обучения системы.В следующих итерациях обучающая выборка может создаваться как направленным поиском в Интернете, так и в ранее созданных базах данных, содержащих скачанные из сети документы.

- (б) Поиск по ключевым словам в открытых веб-источниках с целью создания больших корпусов ЕЯ-текстов (big data). Осуществляется циклическими итерациями. Алгоритм поиска и сбора данных подразделяется на несколько «потоков»:
- i. **Общий мониторинг Интернета.** Направленный семантический поиск и сбор данных из открытых источников по заданным ключевым словам осуществляется авторской системой «KeyCrawler для систем извлечения знаний» [7]. KeyCrawler — оригинальный семантический навигатор по Интернету, обеспечивающий извлечение информационных объектов по исходным семантическим образам — наборам ключевых и опорных терминов. KeyCrawler был успешно применен при разработке авторского метода АП-ПО в проекте РФФИ 13-07-00272. Для построения массива документов была использована модифицированная Keywen-технология (KeyCrawler-2) построения интернет-корпусов, специально настроенная на поиск в Интернете текстов, ассоциированных с заданными ключевыми терминами из выбранных предметных областей. Для создания веб-корпуса необходимо выполнение следующих операций: 1) загрузка больших объемов данных (big data) из Интернета, извлечение текстовой информации; 2) идентификация языка загруженных текстов, удаление «неправильных» документов; 3) сегментирование текста в абзацы и предложения; 4) удаление дублированного контента (идентичные или частично идентичные сегменты текста); 5) токенизация — сегментирование текста на слова; 6) лингвистическая аннотация-лемматизация и тегирование; 7) выгрузка получившегося корпуса в корпус-менеджер (то есть, создание соответствующих индексных структур), которые делают корпус доступными для пользователей. Технология Keywen выполняет все эти операции, включая удаление страниц нецелевого языка (и прочего «шума»), отсеечение HTML-кодов и проверку релевантности. KeyCrawler-2 по заданным ключевым терминам строит не только интернет-корпуса ЕЯ-текстов, но также тематические коллекции научных документов (PDF) с названиями, авторами и библиографическими ссылками.
 - ii. Создание базы данных первоначального неструктурированного на подтемы корпуса ЕЯ-текстов на основе проведенного семантического поиска.
 - iii. Специализированный мониторинг соцсетей. Методом сбора данных из соцсетей выступает система IQBuzz. IQBuzz — это Интернет-сервис отслеживания соцсетей, позволяющий анализировать интенсивность коммуникации по отдельным темам, выгружать массивы данных и др.
- (в) Семантический поиск по ключевым словам-именованным сущностям, автоматически выделенным с помощью разработанного авторским коллективом лингвистического процессора BREF (Book Reference). BREF выделяет сами научные статьи, их параметры (имена авторов, названия статей, локации и др.) и библиографические ссылки (в виде XML-структуры). Процессор BREF (Book Reference), выделяющий из статей формальные данные (название, авторы, библиографические ссылки и т.д.), представляет их в формате XML-файла

с простой и самодокументируемой структурой [8]. Процессор обработки ссылок и списка литературы BREF был реализован на базе лингвистического процессора Pullenti (Puller of Entities, www.pullenti.ru), который под псевдонимом Pink в ходе соревнования FactRuEval на конференции Диалог-2016 занял первые места на большинстве дорожек (см. <http://www.dialog-21.ru/evaluation/2016/ner/>). В 2017 г. BREF тестировался на множестве обучающих выборок, в результате чего его алгоритмы были усовершенствованы.

- (г) Семантический поиск по ключевым словам, автоматически выделенным из первоначальных обучающих корпусов и текстовых коллекций с помощью методов статистики (например, по их частотным характеристикам).
2. Составление частотного словаря наиболее значимых словосочетаний (терминов, ключевых слов, фрагментов текста) для каждой предметной области.
 3. Автоматический отбор ключевых слов из больших полнотекстовых корпусов, значимых для целей тематической кластеризации, с помощью методов тематического моделирования (в частности, LDA).
На основе выделяемых признаков осуществляется кластеризация и тематическое моделирование, в результате которых становится возможным а) автоматизированное разбиение полученного корпуса на тематические коллекции документов и б) расширение первоначально построенной системы тематических категорий при пополнении корпуса новыми текстами.
 4. Автоматизированное разбиение полученных выборок из Интернета и БД big data (в результате различных вышперечисленных видов поиска) на тематические корпуса и коллекции документов определённой предметной области методами статистики, кластеризации и тема-

тического моделирования (в частности, LDA).

5. Обнаружение формально связанных документов путём выявления и подсчёта цитирований с помощью лингвистического процессора BREF, автоматически выделяющего библиографические ссылки из отдельных научных документов и сборников конференций.
6. Автоматизированный подсчёт индекса научного цитирования (ИНЦ). ИНЦ рассчитывается как количество формальных библиографических ссылок на документ (или автора). В данном проекте ИНЦ вычисляется не по базам РИНЦ, Scopus, Web of Science, а по тематическим выборкам научных статей. *Расчёт ИНЦ* осуществляется с помощью процессора BREF, выделяющего из коллекции документов библиографические ссылки в виде направленного графа формальных ссылок. Далее подсчитывается количество формальных ссылок (входящих дуг графа) на каждую статью.
7. Вычисление *семантической меры подобия (меры неявной связанности) между текстами*. Расчёт ведётся на созданных тематических корпусах (коллекциях). Мера подобия — предложенный авторским коллективом показатель — основана на соотношении явных (библиографических, прямое цитирование) и неявных ссылок, когда тексты связаны через общие (или сходные) текстовые фрагменты. Параметры алгоритма расчёта данной меры оптимизируются по максимуму корреляции индексов ИНЦ и ИКНЦ.

В простейшем случае мера неявного подобия (мера сходства) — measure of implicit connection (MIC) — рассчитывается по следующей формуле:

$$MIC = \frac{(S_{intersection})}{\min(S_{text_1}, S_{text_2})} \times 100\%,$$

где: $S_{intersection}$ представляет собой сумму весов всех терминов (текстовых фрагментов), включенных в оба текста;

S_{text_1} — это сумма весов ключевых словосочетаний (терминов) первого текста, S_{text_2} представляет сумму весов ключевых словосочетаний (терминов) второго текста. Мера неявного подобия двух идентичных текстов равна 100%.

Можно считать, что между двумя текстами есть неявная ссылка/связь, если их мера подобия составляет более 8%. В этом случае вероятность наличия связи по библиографическим ссылкам между этими текстами более 70%. Использование данной меры помогает находить пропущенные библиографические ссылки между статьями. Таким образом, наличие неявной ссылки между статьями можно определить с помощью предложенной меры подобия текстов. В экспериментах 2017 г. используются также методы Word2Vec (для определения семантической близости терминов) и метод WMD (для определения семантической близости тем и групп терминов).

8. Выявление имплицитной/неявной связанности документов между собой с помощью лингвистического процессора неявных ссылок (ЛПНС), обнаруживающего неявные ссылки с заданными параметрами (пороги минимального и максимального семантического подобия, учет времени статей, и т. д.). ЛПНС настраивается с помощью машинного обучения так, чтобы корреляция между индексами ИНЦ и индексом контекстного цитирования (ИКНЦ) была максимальной.

В процессоре ЛПНС версии 2017 года для сравнения семантического подобия текстов используется мера близости терминов на основе Word2Vec (технология нейронных сетей), а также мера WMD (Word Mover's Distance), которая хорошо оценивает расстояние между множествами терминов.

9. Автоматизированный расчёт индекса контекстного научного цитирования (ИКНЦ) на основе текущей версии лингвистического процессора неявных ссылок (ЛПНС), работающего по ме-

тоду обнаружения релевантных фраз. ИКНЦ вычисляется на материале текстов из тематических коллекций методом анализа текста статьи и сравнения его с текстами других статей с использованием меры семантического подобия (semantic similarity measure). В одной из возможных реализаций алгоритма вычисления ИКНЦ равен числу неявных связей, параметры которых определяются через семантическое сходство, так что корреляция между ИКНЦ и ИНЦ максимизируется. Далее осуществляется преобразование информации об ИКНЦ в различную форму: по статьям (рейтинги статей и 3D-визуализация статей с учетом значимости/индекса), по авторам, по организациям и т. д.

Мера подобия двух произвольных текстов (семантическое сходство) и ИКНЦ

Мера подобия (мера семантического сходства) — новый авторский показатель. Рассмотрим его подробнее. Экспериментальный корпус из 120 связанных статей по теме КГВ, взятых из доступных в Интернете российских научных публикаций был создан в 2016 г. Первоначально нашлось более 1000 статей, откуда затем были автоматически выбраны 120 статей, связанных формальной библиографической ссылкой. Мера подобия выбранных пар менялась от 1,39% до 99%. Большинство рассматриваемых пар имели меру подобия менее 4% и вероятность наличия связи менее 30%. Было принято, что между двумя текстами есть имплицитная связь, если их мера подобия составляет более 8%. Вероятность наличия связи по библиографическим ссылкам между такими текстами более 70%. В 2017 г. этот корпус был расширен до 250 связанных статей. Самая цитируемая статья в обоих корпусах: А. Бондарев, В. Галлактионов «Анализ многомерных данных в задачах многопараметрической оптимизации с применением методов визуализации» (см. <http://sv-journal.org/2012-2/01/>).

Эта статья цитируется 7 раз в корпусе из 120 статей и 13 раз в корпусе 250 статей.

В статьях обеих коллекций представлено два типа связанности: явная (формальная) связь, когда тексты связаны библиографическими ссылками, и неявная связь, когда тексты связаны через общие (или сходные) текстовые фрагменты. Мера подобия (семантического сходства) основана на соотношении явных и неявных ссылок. На основе эксперимента показано, что оба типа связанности коррелированы. Найдены параметры оптимальной обработки текста, когда корреляция максимальна и достигает около 55%. В простейшем случае вес текстового фрагмента равен длине этого фрагмента, если длина превышает определенный порог. Вес равен нулю, если длина меньше этого порога. На рис.1 показано, что корреляция достигает максимума при пороге = 7. Преимущество такого подхода заключается в том, что он не требует ручной работы экспертов для оценки степени связанности и сходства текстов, так как основывается на автоматизации оценки авторитетного мнения ученых, которое выражается в ссылках и в индексе научного цитирования. Была выполнена первая реализация нахождения индексов цитирования на коллекциях по компьютерной графике (КГВ). В этой реализации ИНЦ равен числу формальных библиографических ссылок, а ИКНЦ рассчитывается как количество неявных ссылок на каждый текст коллекции.

Важнейший параметр неявной связи между двумя текстами — это минимальное значение (пороговое значение) меры подобия этих текстов. Было принято, что две статьи имеют неявную ссылку между собой, если мера их семантического подобия выше этого порога.

Для проведения экспериментов с ИКНЦ использовались два корпуса по компьютерной графике, содержащих 120 и 250 статей, связанных официальными библиографическими ссылками. Результаты эксперимента показаны на рис.1, где вертикальная ось пред-

ставляет собой корреляцию между ИКНЦ и ИНЦ, в то время как горизонтальная ось представляет значение порога семантического подобия в сотых долях процента.

Как видно из рис.1, оптимальный параметр (минимальное пороговое значение семантического сходства) составляет около 0,7% для коллекций из 120 и 250 статей. Другие оптимальные параметры можно найти аналогичным способом. При оптимизации сразу нескольких параметров (например, минимального и максимального порога меры семантического подобия), корреляция между индексами ИКНЦ и ИНЦ достигла 40%. Это обнадеживающий результат, так как оптимальные параметры имеют схожие значения в разных коллекциях, что позволяет расширить объём применения данного метода и использовать его в разных предметных областях. Оптимальные параметры неявных связей, вычисленные при создании коллекций научных статей путем максимизации корреляции между ИКНЦ и ИНЦ, могут применяться для анализа других коллекций и доменов, даже тех, где относительно редко встречаются формальные ссылки (например, в текстах социальных сетей). Таким образом, использование ИКНЦ помогает обнаруживать наиболее значимые статьи и более точно отражать смысловое сходство между ними.

Математическая вероятностная модель зависимости количества прямых цитирований от количества неявных ссылок и их параметров

Рассматриваемая ниже вероятностная регрессионная модель зависимости количества формальных цитирований от количества неявных ссылок и их параметров является еще одной из возможных реализаций алгоритма вычисления ИКНЦ, которая позволяет рассчитать ИКНЦ в тех же единицах, как и ИНЦ, а также позволяет оценить качество и точность модели через коэффициент детерминации. Эта модель рассчитывает предсказанное количество формальных цитирований, которое удобно при-

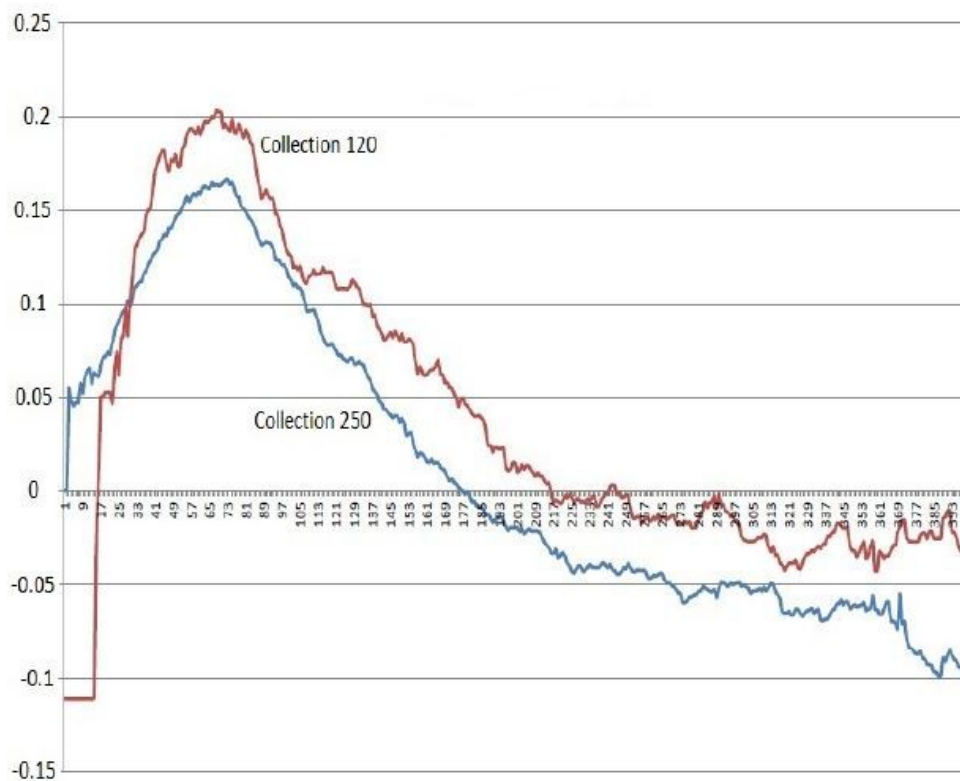


Рис. 1: Корреляция между ИКНЦ и ИНЦ, зависящая от значения порога семантического подобию.

нять за нормализованное значение ИКНЦ и которое более привычно, чем количество неявных ссылок.

Данная прогностическая модель строится на основе моделирования значений независимых переменных и их коэффициентов регрессии. Корреляция и регрессия — связанные между собой и широко используемые методы для определения силы связи между переменными. Корреляция представляет собой меру ассоциации, а регрессия обеспечивает средство прогнозирования одной зависимой переменной от других независимых переменных. Специальный лингвистический процессор ищет неявные ссылки с заданными параметрами. В качестве математической вероятностной модели зависимости количества формальных цитирований от количества неявных ссылок и их параметров используется линейная регрессионная модель.

Регрессия определяется по следующей формуле: $Y = f(X, A) + \textit{epsilon}$; $E(\textit{epsilon}) = 0$, где: A — параметры модели, $\textit{epsilon}$ — случайная ошибка модели. Регрессия называется линейной (множественной), если функция регрессии $f(X, A)$ имеет вид: $f(X, A) = A_0 + A_1 \times X_1 + A_2 \times X_2 + \dots + A_k \times X_k$, где: A_i — коэффициенты регрессии, X_i — объясняющие/влияющие/независимые переменные (факторы модели), k — количество факторов модели. Линейная регрессия моделирует связь между зависимой переменной и независимыми переменными путем подгонки линейной функции регрессии к наблюдаемым данным. Количество неявных ссылок и их параметры считаются независимыми объясняющими переменными, а предсказанное количество формальных цитирований — зависимой переменной.

Линейный коэффициент корреляции (или коэффициент корреляции Пирсона) изменяется от -1 до $+1$ и рассчитывается по

формуле [9]:

$$R(X, Y) = \frac{\text{sum}((X - \text{sred}(X)) \times (Y - \text{sred}(Y)))}{\text{sigma}(X) \times \text{sigma}(Y)}$$

где:

$\text{sigma}(X) = \text{sqrt}(\text{sum}(X - \text{sred}(X)^2))$ – стандартное отклонение ряда/выборки X;

$\text{sigma}(Y) = \text{sqrt}(\text{sum}(Y - \text{sred}(Y)^2))$ – стандартное отклонение ряда/выборки Y;

$\text{sred}(X) = \frac{\text{sum}(X)}{n}$ – среднее арифметическое.

В формальной математической записи формула коэффициента корреляции имеет вид:

$$r_{XY} = \frac{\text{cov}_{XY}}{\sigma_X \sigma_Y} = \frac{\sum(X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum(X - \bar{X})^2 \sum(Y - \bar{Y})^2}}$$

$$\bar{X} = \frac{1}{n} \sum_{t=1}^n X_t \quad \bar{Y} = \frac{1}{n} \sum_{t=1}^n Y_t.$$

В модели линейной регрессии коэффициенты регрессии связаны с соответствующими коэффициентами корреляции. Зная корреляцию, можно рассчитать *коэффициент регрессии* по формуле: $A_i = R(X_i, Y) \times (\text{sigma}(Y)/\text{sigma}(X_i))$; где: $\text{sigma}(X_i)$ – среднеквадратическое отклонение соответствующего факторного признака, Y – количество формальных ссылок. В нашей регрессионной модели значение функции $f(X, A)$ – это предсказанное количество формальных цитирований, а объясняющие переменные X_i – факторы, влияющие на количество цитирований. Такими факторами являются семантическое подобие текстов (semantic similarity), индекс Хирша автора, рейтинг журнала/конференции/источника и т. д.

В нашем исследовании мы концентрируемся на факторе $X1$ – семантического подобия текста рассматриваемой статьи с другими текстами, т. к. это наиболее прямой показатель качества статьи и ее влияния на другие тексты. При таком подходе результаты исследования можно будет широко применять в различных областях, где автор и рейтинг источника часто неизвестны, например, при выявлении значимых

интернет-текстов. Фактор семантического подобия статьи – это некоторая функция, которая рассчитывает количество неявных ссылок на статью в зависимости от текста самой статьи, текстов в анализируемой коллекции, параметров алгоритма расчета неявных ссылок. Формулу фактора подобия $X1$ можно представить в виде некоторой функции следующего вида: $X1 = Fsim$ (текст статьи, тексты коллекции, тексты в Интернете, параметры неявных ссылок).

Параметры неявных ссылок – это параметры алгоритма расчета семантического подобия текстов, такие как минимальный и максимальный пороги семантического подобия, минимальная длина учитываемых терминов, а также параметры, определяющие необходимость использования других подходов для выявления этих ссылок. В дальнейшем предполагается применение следующих методов: Word2Vec, Word Mover's Distance (WMD), лингвистический процессор для выделения объектов (люди, организации, адреса), алгоритмы тематического моделирования, дискриминационные корзины терминов (Discriminative term bucketing) и т. д.

Таким образом, в проекте исследуются различные функции расчета неявных ссылок $Fsim()$, при которых достигается максимум корреляции между фактором $X1 = Fsim()$ и количеством формальных ссылок Y . Чем выше корреляция, тем качественнее регрессионная модель. Для определения качества регрессионной модели используется коэффициент детерминации. Коэффициент детерминации (R^2 – R-квадрат) – это доля дисперсии зависимой переменной, объясняемая рассматриваемой моделью зависимости, то есть объясняющими переменными. Коэффициент детерминации равен квадрату (множественного) коэффициента корреляции. Итак, в данном проекте исследуются различные алгоритмы семантического подобия текстов и поиска неявных ссылок $Fsim()$, при которых достигается максимум корреляции между функцией подобия $Fsim()$ и количеством формальных

ссылки Y . Чем выше корреляция, тем качественнее модель и тем большую долю вариации Y (т. е. индекса ИНЦ) она объясняет.

Заключение

Для выявления связей между научными статьями (явные и неявные ссылки), а также оценки качества научных статей предложены новые авторские показатели — индекс контекстного научного цитирования (ИКНЦ) и мера семантического подобия. Получены текстовые корпуса, в которых выявлены формальные и неявные ссылки между документами. Предложена методика автоматизированного построения ИКНЦ. В ходе экспериментов на ПО КГВ выделен один из параметров алгоритма (минимальная длина анализируемых слов), при котором корреляция между индексом научного цитирования (ИНЦ) и ИКНЦ достигает максимума. Был найден один из важных параметров неявной связи между двумя текстами — это минимальное (пороговое) значение меры подобия этих текстов. Было принято, что две статьи имеют неявную ссылку между собой, если мера их семантического подобия выше этого порога. Алгоритм расчёта ИКНЦ использует меру семантического подобия текстов, настраиваемую по максимуму корреляции между библиографическими и неявными ссылками. Значение найденных параметров подтверждено экспериментально [6].

Описана математическая вероятностная модель зависимости количества прямых цитирований от количества неявных ссылок и их параметров.

В дальнейшем планируется: 1) улучшить оценку ИКНЦ и меры семантического подобия, используя Word2Vec, Doc2Vec, WMD и другие новые методы; 2) улучшить оценку ИКНЦ и меры семантического подобия, учитывая время публикации статей и направления неявных ссылок; 3) проверить полученные результаты в разных коллекциях и в других предметных областях, а также на других языках; 4) построить многоязычный ИКНЦ и рассчиты-

вать неявные ссылки между документами на различных языках, используя базу данных BabelNet.org; 5) создать улучшенную версию ИКНЦ, которая зависит не только от количества неявных ссылок, но также и от их качества.

Литература

- [1] Клименко С.В., Шарнин М.М., Хакимова А.Ф., Золотарев О.В., Мацкевич А.Г. Методы оценки качества и влияния (impact) научных статей для повышения объективности индекса научного цитирования. // «Вестник Российского нового университета». М.: РосНОУ. 2016. Выпуск 3. С. 51-59.
- [2] Шарнин М.М., Сомин Н.В. Об основных типах связанности между текстовыми документами. // Системы и средства информатики, 2017. Т. 27. №1. С. 104-111.
- [3] Демидов А.О., Шарнин М.М. Построение развивающейся во времени системы тематических категорий для пополняемого корпуса текстов. // СРТГ1617 Труды международной научной конференции «Физико-Техническая Информатика СРТ2017». Ларнака, Республика Кипр, 07-14 мая 2017 г. Институт физико-технической информатики (ИФТИ). 2017. С.166-171.
- [4] Шарнин М.М., Клименко С.В., Хакимова А.Х., Сюракшина Ю.В. Методика построения ассоциативно-иерархического портрета предметной области. // СРТГ1617 Труды международной научной конференции «Физико-Техническая Информатика СРТ2017». Ларнака, Республика Кипр, 07-14 мая 2017 г. Институт физико-технической информатики (ИФТИ). 2017. С.343-349.
- [5] Орлов Д.Е., Клименко С.В., Орлова Н.А., Шарнин М.М., Слепцов И.О. Построение и анализ тематических коллекций научных публикаций по тематике «Компьютерная графика, визуализация и виртуальная реальность» // Сборник трудов Международной конференции «Ситуационные центры и информационно-аналитические системы класса 4i для задач мониторинга и безопасности» SCVRT2015-16. Пушино. ЦарьГрад. 21-24 ноября 2015-2016 гг.. С.36-42.
- [6] Michael Charnine, Stanislav Klimenko. Semantic cyberspace of scientific papers. // Proceedings of the 2017 International Conference on Cyberworlds. 20-22 September 2017. Chester. United Kingdom. pp.146-149.
- [7] Charnine M., Charnine V. Keywen category structure. // Wordclay, USA, 2008. 60 p.
- [8] Золотарев О.В., Шарнин М.М., Клименко С.В., Кузнецов К.И. Система PullEnti — извлечение информации из текстов естественного языка и автоматизированное построение информационных систем. // Международная Школа-семинар «Ситуационные центры и информационно-аналитические системы класса 4i для задач мониторинга и безопасности» SCVRT2015-16.

- Пуццино. ЦарьГрад. 21-24 ноября 2015-2016 гг. С.14-21.
- [9] Общая теория статистики: Учебник под ред. Р. А. Шмойловой. 3-е издание, переработанное. // Москва: Финансы и Статистика. 2002.
- [10] Соколов Е.Г., Казанцев А.С., Шарнин М.М. Применение технологии нейронных сетей и векторных операций для выявления объектов различных классов и их связей в большом корпусе текстов. // Труды международной конференции «Ситуационные центры и информационно-аналитические системы класса 4i для задач мониторинга и безопасности (SCVRT2017)». Москва — Протвино. 28-30 ноября 2017. С.131-135.

Определение интеллектуального потенциала и оценка вклада экспертов при использовании системы коллективного согласования решений в сложных случаях анализа текстов*

В. И. Протасов¹, З. Е. Потапова¹, М. М. Шарнин², К. А. Колесникова³
protvlad@gmail.com

¹Московский авиационный институт (национальный исследовательский университет)

Протвино, ²АНО «Институт физико-технической информатики»

³Московский физико-технический институт (государственный университет)

При автоматическом анализе текстов на наличие предмета интереса существует проблема идентификации его наличия в сложных случаях, особенно когда возможны несколько толкований. Предполагается также, что автоматическая система распознавания справляется с анализом большинства текстов и может выделять относительно редкие случаи, где «интеллекта» системы не достаточно для надежной идентификации искомого смысла в тексте. Эти случаи должны передаваться системой сертифицированным экспертам или их группам для дальнейшего анализа. В этом случае возникает две проблемы. Первая заключается в том, что требуется информация об интеллектуальном потенциале экспертов и, соответственно, о методологии получения этой характеристики. Вторая проблема заключается в справедливой оценке вклада каждого эксперта в общую работу.

Ключевые слова: интеллектуальный потенциал, эксперт, система коллективного согласования решений, анализ текстов

Hardware and software tools for augmented reality system for Automated Control Systems*

V. I. Protasov¹, Z. E. Potapova¹, M. M. Sharnin², K. A. Kolesnikova³

¹Moscow Aviation Institute (National Research University)

Protvino, ²ANO "Institute of computing for physics and technology"

³Moscow Institute of Physics and Technology (State University)

In the automatic analysis of texts on the existence of an object of interest, there is the problem of identifying its presence in complex cases, especially when several interpretations are possible. It is also assumed that the automatic recognition system copes with the analysis of most textures and can allocate relatively rare cases where the "intelligence" of the system is not enough to reliably identify the desired meaning in the text. These cases should be transmitted by the system to certified experts or their groups for further analysis. In this case, there are two problems. The first is that it requires information on the intellectual potential of experts and, accordingly, on the methodology for obtaining this characteristic. The second problem is a fair assessment of the contribution of each expert to the overall work.

Keywords: intellectual potential, expert, system of collective agreement of decisions, analysis of texts.

Введение

В нашей предыдущей работе [1] рассматривался вопрос об уменьшении вероятности неправильного решения группы экспертов и было доказано, что это возможно только при выполнении ряда условий. Одним из таких условий и является не превышение каждым экспертом определенной величины определенной в этой работе величины ошибки эксперта. Ясно, что при соблюдении этого условия необходимым для

гарантированного решения сложных случаев группой экспертов является наличие достаточного интеллектуального потенциала у каждого эксперта. Как было показано в [1], для успешного решения сложных случаев анализа текстов группами экспертов наиболее подходящим является метод эволюционного согласования решений [2].

Метод эволюционного согласования решений

Метод устроен следующим образом (рис. 1). На первой стадии — стадии формирования популяции решений эксперты, работая поодиночке, генерируют возможные ответы,

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 16-07-00756, 16-29-09527, 18-07-20030, 18-37-10008.

причем им предоставляется возможность отказаться от ответа. На второй и последующих стадиях — стадиях обмена вариантами эксперты, которые отказались от ответа, выбирают из предложенных ответов правильные, на их взгляд, ответы или вновь отказываются от ответа. Процесс итераций продолжается до тех пор, когда у большинства экспертов получится одинаковый ответ. Ясно, что этот ответ может быть трех видов: правильный, неправильный или «не могу решить». Величины вероятностей таких ответов зависят как от числа экспертов в группе M , трудности задачи θ так и от исходных вероятностей генерации правильных и неправильных ответов у одиночных экспертов (G_R и G_N соответственно) и от вероятностей правильного или неправильного выбора чужих ответов на стадиях согласования решений (E_R и E_N). В [1] доказано, что при использовании этого метода и соблюдении ряда условий, вероятность правильного решения группы экспертов, по сравнению с вероятностью такого решения одиночным экспертом, может достигать единицы, а вероятность неправильного решения снизиться до нуля.

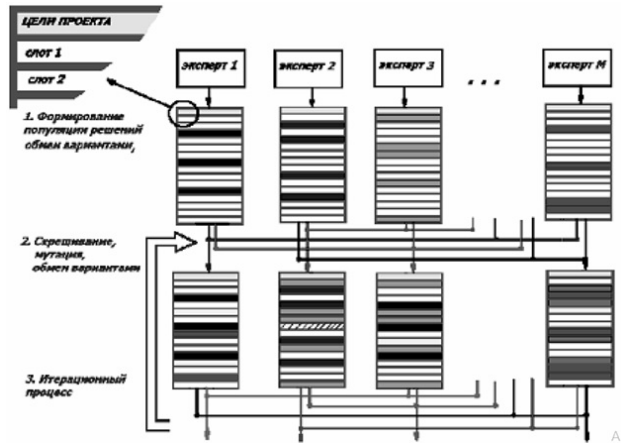


Рис. 1: Схема МЭС

Как легко показать, в случае сильных экспертов метод можно использовать при одной итерации согласования и достаточном числе экспертов.

После заполнения на первом такте малого числа слотов какого-либо проекта каждым

из экспертов, они во втором такте по заранее заданному регулярному графу связей будут получать чужие решения и в меру своей, компетентности по оценке чужих решений, будут наполнять свои слоты, преодолевая барьер в 0.5. Далее, в третьем такте, сработает эффект Кондорсе [3] и при голосовании по каждому слоту будет получено полное решение. Для выхода на полное решение необходимым и достаточным будет подбор соответствующих значений M для данных G_R, G_N, E_R и E_N , характеризующих соответствующие компетентности интеллектуальных агентов.

Разумеется, при генерации нулевого приближения необходимо обеспечить вероятностный характер заполнения слотов. Регулярный граф связей можно применять в силу того, что мы проводим только одну итерацию согласования решений и случайный подбор вариантов, необходимый в случае применения генетических алгоритмов в нашем случае не нужен. Следует отметить также, что каждый из трех тактов выполняется всеми агентами параллельно и общее время работы не зависит от их количества.

Принципиально важными полученными результатами являются и для конструирования искусственного интеллекта из большого количества однотипных модулей, например, ансамблей нейронов или простейших компьютеров, образующих однородную среду и обладающих «низким» интеллектом. На рис. 2 приведена схема такого 3-х тактного усилителя интеллекта [4].



Рис. 2: 3-х тактный усилитель интеллекта

Определение интеллектуального потенциала эксперта

При формировании группы экспертов для решения конкретной задачи важными являются вопросы подбора экспертов исходя из их креативных способностей, как генераторов идей, так и оценителей чужих решений. Если есть соответствующая статистика решения подобных задач различными экспертами с измеренными значениями креативных способностей экспертов, то применение компьютерной модели позволяет оценить для данного состава экспертов соответствующие значения вероятностей правильных или неправильных решений, а также вероятностей ответов «не знаю».

В принципе, способности экспертов также могут быть оценены путем индивидуального тестирования с использованием специально подобранных тестов с известными ответами подобно тому, как измеряются значения IQ в тестах Айзенка [5]. Такие тесты необходимо разбить на две группы. В первой группе вопросы подбираются таким образом, чтобы можно было оценить способности эксперта к генерации идей. Правильный ответ при этом формулируется экспертом исходя из своей компетенции. Во второй группе тестов на каждый вопрос предлагается несколько возможных ответов, один из которых правильный и именно его и нужно выбрать эксперту. Так оцениваются способности эксперта к оцениванию чужих решений.

В том и другом случае открытым остается следующий вопрос, какой эксперт может дать больший вклад в общую работу, тот у которого выше значение G_R при меньшем значении E_R или наоборот? Также можно поставить этот вопрос следующим образом. Допустим, у эксперта A известны значения креативных способностей G_R и E_R , а у эксперта B известно только одна из этих величин. Спрашивается, какой величины должна быть вторая характеристика, чтобы интеллектуальный агент B давал такой же

вклад в общее решение, как и агент A , или же как доказать, что для данной тройки чисел это в принципе невозможно. Таким образом, мы приходим к понятию интеллектуального потенциала эксперта, то есть такой его обобщенной характеристики, которая позволяет однозначно определить его индивидуальный рейтинг для работы в системах коллективного интеллекта.

Проблема решается следующим образом. Исходя из компьютерной модели МЭС для данной пары значений G_R и E_R i -го эксперта, мы можем подобрать такое значение M – числа экспертов группы с такими же креативными характеристиками, чтобы этот коллектив за одну итерацию получил полное решение. Величину, пропорциональную обратному значению M , определяющую вклад данного эксперта в общее решение и будем называть интеллектуальным потенциалом IP данного эксперта.

Для удобства в качестве коэффициента пропорциональности можно взять 1000 условных единиц и округлить до целого числа получившееся значение интеллектуального потенциала

$$IP = \left\lfloor \frac{1000}{M} \right\rfloor. \quad (1)$$

Поскольку при $G_R \gg G_N$ выражение для оценки сверху вероятности правильного решения Q группы из M экспертов можно записать в виде

$$Q = 1 - (1 - G_R)^M, \quad (2)$$

то очевидно, что небольшой линейный прирост G_R у экспертов приводит к существенному нелинейному росту их интеллектуального потенциала, причем его отличие у разных экспертов может расти по экспоненте и достигать нескольких порядков.

Обработка результатов множества компьютерных экспериментов с использованием метода наименьших квадратов с различными парами значений G_R и E_R (при $G_R \gg G_N$ и $E_R \gg E_N$) дала следующую эмпири-

ческую формулу:

$$IP = \left[\frac{1000}{\frac{1.2075}{G_R E_R} + \frac{5.675}{G_R} \frac{1.148}{E_R} - 2.465} \right]. \quad (3)$$

Максимальная относительная погрешность между значениями величин IP , полученных с помощью компьютерной модели и расчетами по этой формуле не превысила 0.005 при относительной погрешности 0.037.

Исходя из полученных результатов, можно представить следующую процедуру определения величины IP у экспертов. Из известных результатов работы по анализу текстов в сложных случаях, полученных ранее другими экспертами с высокой квалификацией, составляются тестовые вопросы с известными жюри ответами. Тестовые вопросы с закрытыми ответами служат для определения величин G_R и G_N , а тестовые вопросы с вариантами ответов — для определения величин E_R и E_N . Эксперты, у которых не выполняются величин G_R и G_N , а тестовые вопросы с вариантами ответов — для определения величин E_R и E_N к работе для анализа сложных текстов не допускаются, а у допущенных, исходя из (3), определяются величины IP .

В дальнейшем целесообразно комплектовать группы из экспертов с примерно равными интеллектуальными потенциалами.

Оценка вклада экспертов

В практике оценки вклада каждого участника проекта в общий продукт существует проблема справедливой оценки качества и количества затраченного интеллектуального труда. Для исключения субъективного фактора, искажающего такую оценку, необходимо построить автоматическую процедуру этой оценки, основанной на математически выверенных принципах и не вызывающей споров участников. В общем случае задачи компьютерной лингвистики решаются, как правило, переменным числом участников разной квалификации и проблема выявления вклада каждого участника в

зависимости от его квалификации и затраченного им времени не является тривиальной.

Рассмотрим несколько случаев. Начнем с самого простого. Имеется длительный проект с большим количеством заданий (слотов) и все M экспертов затрачивают одинаковое время, выполнив примерно одинаковое количество «квантов» работы. За квант работы принимается генерация одного слота проекта или один акт экспертизы чужих решений тоже одного слота. Предполагается, что эксперты были подвергнуты сертификации на подобных задачах и их ошибаемость не превышает полученной в [1] критической величины. В таком случае, как показано в [1], в работе над проектом будут получены либо правильные заполнения слотов с пренебрежимо малой вероятностью ошибки, либо ответы типа «ответ неизвестен». Допустим, что все кванты работы всех участников проекта запоминаются программой, координирующей работу экспертов. Исходя из истории работы, можно определить величины G_R, G_N, E_R и E_N у всех экспертов и при выполнении неравенств $G_R \gg G_N$ и $E_R \gg E_N$ по формуле (3) оценить интеллектуальный потенциал IP каждого эксперта. Вклад каждого из них в этом случае считать пропорциональным этому потенциалу, то есть можно определить вклад W каждого актора как отношение его IP к сумме IP всех акторов (стоимость проекта принята за единицу):

$$W_i^{(1)} = \frac{IP_i}{\sum_{i=1}^M IP_i}. \quad (4)$$

В случае, если неравенства не выполняются, и вероятности ошибочных решений соизмеримы с вероятностями правильных решений, выполняются монте-карловские вероятностные расчеты на большой статистике с использованием концепции виртуальных экспертов, обладающих такими же величинами G_R, G_N, E_R и E_N , как и у реальных экспертов по методике, изложенной в [6].

Следующий простой случай оценки вклада экспертов в консолидированный продукт можно представить, когда проект состоит из одного слота и невозможно статистически точно оценить интеллектуальный потенциал экспертов IP . В этом случае вклад W можно оценить по величинам IP , измеренных при предварительной сертификации экспертов с использованием тестов той же степени трудности.

Но так бывает не всегда — чаще слоты проекта заполняются не постоянным коллективом актором, а переменным. В этом случае является разумным комбинировать оба подхода. Бывает, что какой-либо актер может выступить как генератор идеи, которая будет подхвачена сообществом, а он рано уйдет из проекта. Так как же учитывать его вклад? А как учитывать вклады других, мало поработавших актеров над проектом, и в силу этого не хватает статистики, чтобы достоверно оценить их вклад? Как и в рассмотренном выше первом случае, по результатам проекта вычислить IP только у актеров, для которых хватает статистики, и они произвели примерно одинаковое количество квантов работы. У мало поработавших экспертов нужно взять их IP по результатам тестирования при сертификации. Для качественной и количественной оценки вклада W каждого эксперта нужно также присвоить ему весовой коэффициент, равный числу N его правильно сгенерированных и оцененных слотов:

$$W_i = \frac{IP_i}{\sum_{i=1}^M IP_i} \times \frac{N_i}{\sum_{i=1}^M N_i}. \quad (5)$$

Здесь следует отметить, что при измерении IP по малому числу тестов (меньше ста) нужно определять погрешности измерения. В дальнейшем у экспертов, мало работающих над проектом нужно брать величину IP , соответствующей нижнему значению $IP - \Delta IP$, где ΔIP — абсолютная статистическая ошибка измерения.

Таблица 1: Результаты работы групп экспертов

i	M	N_R	K	Q	IQ	IP_{\min}	IP_{\max}
1	7	17	0.055	0.34	132	53	200
2	9	21	0.074	0.42	144	71	250
3	11	18	0.061	0.36	135	63	200
4	5	11	0.032	0.22	118	29	125
5	8	14	0.018	0.28	125	48	167
6	12	23	0.041	0.46	146	34	333
7	6	12	0.052	0.24	121	29	143
3-й этап	7	50	0.612	1.0	215	167	333

Результаты экспериментов

По предложенной методике был организован эксперимент с использованием тестов с заранее известными ответами для нескольких групп экспертов. Была поставлена задача по рассчитанным с использованием данной методики величинам IP составить сводную группу лидеров, которая завершит работу отдельных групп, если они не смогут полностью решить тестовые задания.

В таблице 1 приведены результаты этой работы. В первой графе приведен номер группы i , во второй — число экспертов в группе M , в третьей — число правильно заполненных слотов группой на втором этапе N_R , в четвертой — относительная величина заполнения слотов по Кондорсе на первом этапе K , в пятой — относительная величина заполнения слотов после второго этапа Q , в шестой — величина коэффициента интеллекта IQ , рассчитанная по методу Айзенка [5], в 7-й и 8-й измеренные величины IP аутсайдера и лидера группы. Всего экспертам было представлено 50 вариантов для анализа текстов.

На завершающем, третьем этапе, из всех экспертов была выделена группа лидеров, обладающих наилучшими показателями по оценке чужих решений. Они обменялись своими вариантами и дополнили их. После проведения этапа согласования в окончательном варианте решения были выделены слоты, имеющиеся у большей части экспертов этой группы. В последней строке таблицы 1 представлены эти результаты. В завершающей стадии эксперимента сводная группа лучших экспертов получила полное решение — были получены от-

веты на все вопросы теста. Оценка вклада экспертов в этом эксперименте не учитывалась, поскольку эксперты работали бесплатно.

Анализ результатов показывает, что на первом этапе происходит формирование популяции правильно заполненных слотов. Естественно, «суммирование по Кондорсе» на этом этапе дает величину Q , близкую к 0 для данного состава экспертов, поскольку из 3-й графы таблицы 1 видно, что у всех экспертов $G_R < 0.5$. После второго этапа у многих экспертов наблюдается значительный рост числа правильно заполненных слотов, но «суммирование по Кондорсе» все еще дает низкую величину Q относительной доли правильно заполненных слотов (см. 5-ю графу табл. 1). И только на завершающем этапе группа лидеров получила результат $Q = 1.00$.

Заключение

Полученные результаты имеют принципиальное значение для применения метода эволюционного согласования для анализа сложных текстов в случае невысокой квалификации экспертов. При выделении лидеров по величинам IP , измеренным по результатам работы на этапе согласования, метод позволяет получить полное решение сложной проблемы. Можно сделать вывод, что при совместной работе экспертов с использованием метода эволюционного согласования наблюдается синергетический эффект усиления интеллекта. Также можно сделать вывод о возможности получения практически безошибочных решений при использовании групп сертифицированных экспертов. Если квалификации экспертов недостаточно для правильного анализа сложных текстов, то предложенная процедура коллективного согласования решений обеспечивает передачу заданий группам экспертов с более высокой квалификацией.

Работа выполнена по грантам РФФИ № 16-07-00756 «Исследование и разработка семантических методов построения "Индек-

са контекстного научного цитирования"» и № 16-29-09527 «Исследования и разработка методов тематического моделирования для мониторинга, прогноза и визуализации террористической активности в информационном поле Интернет с использованием виртуального окружения».

Литература

- [1] Протасов В.И., Шаронов А.В., Шарнин М.М., Клименко А.С. Минимизация вероятности ошибок в распознавании объектов бортовыми компьютерными системами беспилотных летательных аппаратов// Труды МАИ, 2017, № 92. http://mai.ru/upload/iblock/277/protasov_sharonov_sharnin_klimenko_rus.pdf
- [2] Протасов В.И. Теоретическое обоснование метода эволюционного согласования решений. Материалы II Международной конференции «АИУС 2011». Нальчик, 2011, т. 1, с. 85-99.
- [3] Condorcet, marquis Marie-Jean-Antoine-Nicolas de Caritat. Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. Imprimerie Royale, Paris, 1785.
- [4] Протасов В.И. Метод эволюционного согласования решений — теоретическая основа краудсорсинга. Труды XVIII международной конференции «Технологии будущего для человечества», Ларнака, Кипр, 2013. — Изд. Института физико-технической информатики, с. 15-18.
- [5] IQ-тесты. Издательство «Одиссей» 2008 г.
- [6] Протасов В.И., Колесникова К.А., Карелина Е.В. Иерархические самоуправляемые системы коллективного интеллекта. Труды Международной конференции «Физико-техническая информатика СРТ2017». Кипр, Ларнака, 8-14 мая 2017 г. — Изд. ИФТИ, 2017, с. 318-323.

Индекс идеологического влияния: базовые модули комплексной методологии*

С. В. Клименко¹, М. М. Шарнин², И. В. Галина², А. Г. Мацкевич³,
А. Х. Хакимова¹, Е. Г. Соколов⁴

mc@keywen.com | irn_gl@mail.ru | xmag@mail.ru | aida_khatif@mail.ru | evgeny.sokolov@phystech.edu

Протвино, ¹АНО «Институт физико-технической информатики»

Москва, ²Федеральный исследовательский центр «Информатика и управление» РАН

³Московский технический университет связи и информатики (МТУСИ)

⁴Московский физико-технический институт (Государственный университет)

В статье изложены теоретические и практические аспекты автоматизации методов тематического моделирования (ТМ) для мониторинга, прогноза и визуализации экстремистской активности в Интернете. Проект осуществляется на большом объеме данных (Big Data). Рассмотрены базовые элементы междисциплинарной технологии (математика, лингвистика и др.), реализуемой как единый процесс от сбора данных до получения результата. Технология включает автоматизированное построение нового показателя степени экстремизма Сайта — Индекса Идеологического Влияния (ИИВ), рассчитываемого по неявным ссылкам между документами. Значимость достигнутых результатов заключается в создании автоматизированного инструмента, умеющего не только выявлять группы экстремистской направленности, но и прогнозировать их формирование в Интернете. Оптимальные параметры алгоритма расчёта неявных ссылок находятся автоматически по максимуму корреляции между явными и неявными ссылками.

Ключевые слова: Тематическое моделирование, автоматизированные системы, Big Data, индекс идеологического влияния, экстремизм, Интернет, неявные ссылки.

Index of ideological impact: basic modules of integrated methodology*

S. V. Klimenko¹, M. M. Sharnin², I. V. Galina², A. G. Matskevich¹,
A. K. Khakimova¹, E. G. Sokolov⁶

Protvino, ¹ANO "Institute of computing for physics and technology"

Moscow, ²FRC CSC RAS

³Moscow Technical University of Communication and Informatics (MTUCI)

⁴Moscow Institute of Physics and Technology (State University)

The article describes theoretical and practical aspects of automation of thematic modeling (TM) methods for monitoring, forecasting and visualization of extremist activity on the Internet. The project is implemented on a Big Data. The basic elements of interdisciplinary technology (mathematics, linguistics, etc.), realized as a single process from data collection to obtaining the result, are considered. The technology includes the automated construction of a new indicator of the degree of extremism of the Site — the Index of Ideological Impact calculated by implicit references between documents. The significance of the results achieved is to create an automated tool that can not only identify extremist groups, but also predict their formation on the Internet. The optimal parameters of the algorithm for calculating implicit references are automatically based on the maximum correlation between explicit and implicit references.

Keywords: Thematic modeling, automated systems, Big Data, index of ideological impact, extremism, Internet, implicit references.

1. Введение. Постановка задачи

Автоматизация процесса обнаружения экстремистской активности в сети — нетривиальная задача. Авторский коллектив ряд лет разрабатывает методы тематического моделирования (ТМ) для анализа сете-

вой экстремистской активности (веб-сайты, блоги и аккаунты в социальных сетях, далее — Сайты) [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22].

Цель проекта: разработка методов для мониторинга, прогноза и визуализации экстремистской активности путем анализа текстовой информации в Интернет. Вводится новый показатель степени экстремиз-

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 16-29-09527, 16-07-00756, 18-07-00909 и 18-07-01111, 18-07-20030, 18-37-10008.

ма сайтов и блогов — ИИВ (*индекс идеологического влияния*), обладающий прогнозируемыми свойствами. Расчёт ИИВ основан на вероятностной модели влияния (impact) тем и идей Сайта на темы и идеи Сайтов схожей тематики. Считается, что темы и идеи могут быть адекватно выражены множеством сходных по смыслу лексических фрагментов (близкий аналог: метод скрытого распределения Дирихле — LDA). На основе подобного представления «идей», а также путем их наблюдения во временной динамике, можно идентифицировать вклад каждого Сайта по отношению к выявленным темам, что позволяет обнаружить скрытые связи между различными сайтами/авторами.

Индекс идеологического влияния может быть использован для измерения степени семантического подобия разноязыковых текстов на основании выявленных межъязыковых семантических неявных связей. Межъязыковое определение семантического текстового подобия является важным шагом для обнаружения и оценки межъязыковой миграции идей, в том числе, научных. Исследования в направлении обнаружения межъязыкового плагиата редки. Многоязычный Индекс Идеологического Влияния (ИИВ), равный количеству похожих последующих статей/документов, позволит найти значимые документы и источники происхождения идей независимо от языка. Разрабатываемая методика позволяет выявлять тексты-первоисточники различных идей в социальных сетях и в Интернете, позволяет проследить потоки распространения идей, а также быстрее находить достоверную и качественную информацию из первоисточников. Дополнительно проект затрагивает следующие задачи: мониторинг новых идей и оценка их перспективности по частоте упоминания в различные периоды времени; анализ преемственности идей в многоязычном пространстве.

Конечным итогом работы является создание комплекса программного обеспечения,

включая: 1) расчет обладающего прогнозными свойствами Индекса Идеологического Влияния (ИИВ) для коллекции документов, связанных с предметной областью экстремизма, на основе выявления неявных контекстных ссылок в текстах коллекции с помощью специального лингвистического процессора; 2) применение сервиса мониторинга социальных медиа IQBuzz для анализа предметной области экстремизма в социальных сетях; 3) направленный поиск в Интернете с целью формирования корпуса текстов данной предметной области (ПО); 4) демонстрация возможностей метода неявного контекстного цитирования, включая способность делить на группы и ранжировать по степени идеологического влияния новые тексты/Сайты; 5) визуализация динамики экстремистской активности.

В настоящее время научный коллектив находится на первом этапе создания разрабатываемой технологии.

2. Краткий обзор современных подходов к проблеме

Экстремистская деятельность в сети сложна для исследования, так как фигуранты стараются действовать скрытно (закрытые Сайты, чаты, мессенджеры и т.п.); объём текстов постоянно растёт. Тематический анализ данной ПО плохо поддаётся автоматизации, так как информация, распространяемая экстремистскими группами, разнородна, разноязыка, а короткие сообщения содержат жаргонизмы, намеренно «испорченные» и закодированные слова, что чрезвычайно осложняет семантический анализ. Наиболее часто для анализа используют метод скрытого распределения Дирихле (Latent Dirichlet Allocation, LDA). Методики поиска по ключевым словам или фразам здесь не работают, ибо не позволяют разграничить экстремистские группы от новостных статей. На сегодняшний день наиболее популярные методы автоматизации, вычисляющие релевантность фрагментов текста, основаны на тематиче-

ском моделировании текстов с использованием латентно-семантического анализа. Латентно-семантический анализ работает с матричным представлением коллекции текстов, получаемым с помощью модели «мешок слов» (англ. “bag-of-words”).

2.1. Зарубежные исследования

Существует большой пласт исследований, посвященных связи Сайтов и социальных сетей с экстремизмом. Э. Феррара из университета Южной Каролины и др. [23] использовали данные Twitter, методы краудсорсинга и машинного обучения (логическую регрессию, Random Forests, SVM, Stochastic Gradient Descent) для обнаружения экстремистских групп и прогнозирования их читательской аудитории. Анализируемые пользовательские аккаунты характеризовались с помощью 52 формальных атрибутов, в том числе хэштеги и временные параметры сообщений. Н. Джонсон и др. [24] в 2016 г. разработали математическую модель для изучения закономерностей поведения экстремистских групп в социальной сети «ВКонтакте». При анализе групп использовались хэштеги: #isn, #khilafah, #fisyria, #игиш, #дауля, #халифат и другие.

А. Торнберг и др. [25] исследовали представление ислама на примере крупнейшего интернет-форума в Швеции Flashback (> миллиона пользователей). Анализ проводился с использованием методов корпусной лингвистики, с учетом лингвистического контекста по ключевым терминам, критического дискурс-анализа (CDA) и тематического моделирования (ТМ). Подход, объединяющий CDA и ТМ, по мнению авторов, позволяет выявить лингвистические закономерности во многих корпусах. Анализ показал, что мусульмане изображаются на форуме как однородная внешняя группа, вовлеченная в конфликт, насилие и экстремизм. Исследование периода в несколько лет позволило выявить как продолжающиеся непрерывные темы, так и новые, появившиеся в конкретном году.

Группа исследователей из университета Шри Ланки [26] предложила методологию Advanced Terrorist Detection System (ATDS), изучающую типичное поведение (профиль) террористов и посещаемые ими веб-сайты, с применением семантических векторных пространств и меры семантического подобия текстов. Предлагаемый подход эффективен для выявления деятельности, связанной с терроризмом, в режиме реального времени.

В [27] анализируется качество работы алгоритма LDA для обработки, поиска и классификации идей в рамках системы учета идей (idea management systems).

В [28] выделяются две основные категории методов статистической семантики, основанные на анализе совместного вхождения терминов (co-occurrence) в корпус текстов, которые позволяют построить небольшие плотные вектора для каждого термина. Первая категория — это методы ТМ (LDA, LSI), анализирующие глобальное распределение терминов в корпусе. Вторая — это методы предсказания локального контекста и включения (embedding) слов и документов в семантическое векторное пространство, такие как Word2Vec [29] и Doc2Vec [30]. Обе категории методов имеют свои преимущества и могут быть использованы совместно. Обе позволяют построить семантическое векторное пространство (СВП) и рассчитывать семантические расстояния между терминами и между документами. При этом метод Doc2Vec (paragraph vectors) является расширением метода Word2Vec и позволяет встраивать документы в то же векторное пространство, в котором находятся термины.

В [31] обсуждается мера семантического подобия между запросами и документами, основанная на методе Word Mover’s Distance (WMD), разработанном М. Дж. Кушнер в 2015 году. WMD использует Doc2Vec и позволяет более точно рассчитывать семантические расстояния между короткими фразами и предложениями. На международной конференции “International Workshop on

Semantic Evaluation SemEval-2016” (США) для различных задач, связанных с оценкой семантического подобия, использовался метод Word2Vec в работах [32, 33], а также метод WMD в [34].

2.2. Отечественные исследования

Меры спецслужб в сфере противодействия терроризму подробно освещены В. А. Середневым [35], С. А. Воронцовым [36], С. А. Чуриловым [37]. Отметим успешный открытый проект НЦПТИ (Национальный Центр информационного противодействия терроризму и экстремизму в образовательной среде и сети Интернет). На сайте <http://нципти.рф> можно сообщить о противоправном контенте. Специалисты НЦПТИ в 2015-2016 гг. апробировали следующие методы: выявление и удаление уже размещенных материалов в социальных сетях и других ресурсах Интернет; взаимодействие с общественностью (краудсорсинг).

А. М. Горбачев [38] предлагает использовать работу с Big Data для создания прогностических моделей поведения личности и возможного развития ситуации. «Цифровой след», оставляемый пользователями соцсетей, может быть использован для поведенческого анализа, включая анализ особенностей текста. В исследовании Д. С. Рязанова и др. [39] осуществлён пространственно-временной анализ экстремистской активности. Авторы отметили внутреннюю гетерогенность современных исламистских организаций (движений), использующих террористические методы. Е. Г. Кирченко [40] отмечает, что одним из способов работы ИГИЛ в Twitter является продвижение необходимых хештегов; за неделю появляются около 40 000 твитов в поддержку ИГИЛ.

В работе М. А. Басараб [41] и др. перечисляются требования к системе поддержки принятия решений при выявлении преступных сообществ в сетях, в том числе: 1) мультиплатформенность (включая Windows и Linux); 2) возможность сбора (включая прямой импорт данных из соци-

альной сети), хранения и экспорта данных «больших» соцсетей; 3) развитие возможностей визуализации сетевых графов.

Русскоязычная сеть «ВКонтакте» стала объектом диссертационного исследования С. И. Суслова [42] с целью изучения особенностей поведения самоорганизующихся агрегатов как сообществ в социальных сетях и аудитории этих сообществ. В [42] описано применение алгоритмов кластеризации графов или нахождения сообществ для классификации узлов на основании связей при сетевом анализе. Автор использовал два алгоритма кластеризации: распространения меток и кластеризация методом глухого телефона. Изучение структуры коммуникативного пространства дало возможность разделить всю генеральную совокупность участников на множества в зависимости от особенностей виртуального социального поведения.

Работы М. И. Ананьевой и др. носят обзорный характер [43, 44]: рассматриваются методы, используемые для обнаружения экстремистских текстов из Интернета, и способы анализа этих текстов. Авторами сформулированы следующие проблемы: отсутствуют применимые для автоматического анализа критерии экстремистской направленности текстов; методики опираются в основном на словарные системы, что требует постоянной актуализации словарей, баз данных и правил; используется преимущественно бинарная классификация или разделение на три класса, что не позволяет полностью выявить тематическую и эмоциональную направленность текстов на естественных языках.

Области приложения тематического моделирования (ТМ) подробно описаны К. В. Воронцовым [46]. Метод LDA — классический метод ТМ, который, согласно Воронцову, порождает две проблемы: априорные распределения Дирихле и их обобщения не моделируют явления языка, а применяются только ради математического удобства; сложно совместить большое число функциональных требований в одной модели. Теория аддитив-

ной регуляризации тематических моделей (ARTM) решает обе проблемы, при этом за основу берётся классическая модель вероятностного латентного семантического анализа PLSA. В. В. Савельев [47] применил метод LDA на большом корпусе текстов из ПО профессиональных стандартов, разработанных Министерством труда и социальной защиты. Общее количество стандартов в корпусе составило 801.

А. О. Янина и др. [48] использовали методику ТМ на данных блога habrahabr.ru. Тематическая модель строилась с помощью библиотеки с открытым кодом BigARTM, которая позволяет оптимизировать одновременно несколько критериев качества и находить сжатые векторные тематические представления статей и запросов. Для подбора коэффициентов регуляризации использована «жадная» стратегия последовательного добавления регуляризаторов в тематическую модель. Тематический поиск реализуется путём сравнения тематических векторов запроса и статей по косинусной мере близости.

И. В. Машечкин и др. [49, 50] рассматривают ТМ и методы машинного обучения, направленные на обнаружение потенциально опасной информации в сети. В работе предложен оригинальный метод автоматического аннотирования и выявления ключевых слов для поиска в потоках текстовых сообщений, основанный на ТМ и LDA. Объектом исследования послужил набор эталонных данных (600 Гб): “kavkazchat”, подготовленный в лаборатории Искусственного интеллекта Аризонского университета США. Показано, что метод позволяет: получать релевантные аннотации в виде выдержки наиболее важных предложений из исходного текста; генерировать по тексту релевантные ключевые слова, которые могут быть использованы для поиска соответствующей информации в сети.

3. Базовые элементы технологии

Рассмотрим подробнее элементы авторской автоматизированной технологии. Основные

её модули следующие: 1) алгоритм создания динамически пополняемой лексической БД экстремистских текстов из Интернет, включая социальные сети; 2) сама лексическая БД, служащая для выявления экстремистских сообществ и связей между группами и отдельными людьми; 3) машинное обучение и краудсорсинг, методика автоматизированного определения экстремистских документов с помощью лексического анализа; 4) методика автоматического расчёта меры семантического подобия текстов, с помощью которой определяется наличие неявной ссылки между текстами и которая позволяет группировать Сайты определенной направленности; 5) лингвистический процессор неявных ссылок (ЛПНС), работающий по методу релевантных фраз, для выявления явных и неявных связей между текстами, Сайтами и отдельными аккаунтами; 6) методика автоматического расчёта индекса ИИВ, основанного на анализе неявных связей с помощью лексической БД, который позволяет ранжировать Сайты по степени влияния (impact); то есть — создание программы расчета ИИВ; основанный на ИИВ прогнозный анализ формирования и роста экстремистских групп; 7) методы визуального анализа экстремистской активности, в том числе с целью идентифицировать тревожные события, предшествующие терактам.

Существует множество работ, анализирующих влияние идей, которые основываются на анализе явных и неявных ссылок между статьями [51, 53]. Аналогичные подходы применимы и к анализу экстремистских сайтов [1, 2, 12, 13, 18]. «Идея» семантически представляется в виде множества сходных по смыслу фраз или множества терминов (выявляется методами ТМ и дистрибутивной семантики). Путем извлечения из текстов таких идей, а также путем наблюдения идей во временной динамике можно идентифицировать вклад каждого Сайта по отношению к выявленным темам, что позволяет обнаружить скрытые связи между сайтами/авторами. Лексическая ба-

за выступает в качестве источника выявления экстремистских Сайтов и обнаружения связей (явных и неявных ссылок) между ними. На основании выявленных сетевых связей строится индекс идеологического влияния Сайта (ИИВ), базирующийся на вероятностной модели влияния (impact) идей/фраз некоего Сайта на идеи и фразы других Сайтов схожей тематики. *Введение ИИВ основано на следующей гипотезе: рост идеологического влияния группы в Интернете способствует росту численности этой группы; влияние измеряется количеством новых идей группы, которые получили распространение, а рост численности — количеством дополнительных Сайтов этой группы. ИИВ же рассчитывается автоматически по неявным контекстным ссылкам на Сайт. ИИВ обладает прогнозными свойствами и позволяет делить новые Сайты на группы и ранжировать их по степени влияния.*

3.1. Направленный семантический поиск

Алгоритм поиска и сбора данных подразделяется на 2 «потока» — общий мониторинг Интернета и специализированный мониторинг соцсетей. Методом сбора данных из соцсетей выступает система IQBuzz [54]. Сбор данных из Интернет осуществляется автоскрочкой системой «KeyCrawler Интернета для систем извлечения знаний» [21, 22, 55, 56]. В поисковый запрос включаются термины-антагонисты рассматриваемой экстремистской группы.

IQBuzz — это Интернет-сервис мониторинга соцсетей, позволяющий анализировать интенсивность коммуникации по отдельным темам, определять геолокацию пользователей, фиксировать модальность документа, выгружать массивы данных. IQBuzz предоставляет данные об истории сообщений в различных соцсетях и блогах. KeyCrawler — оригинальный семантический навигатор по Интернет, обеспечивающий извлечение информационных объектов по исходным семантическим образам

— наборам ключевых и опорных терминов. KeyCrawler был успешно применен при разработке метода АППО (ассоциативного портрета предметной области) в проекте РФФИ 13-07-00272.

3.2. База данных (БД) и тезаурус идей

Для хранения коллекции текстов используется БД, которая также хранит списки тем, словари терминов, значимые словосочетания (ЗС) и ключевые фразы, составляющие в совокупности лексические ресурсы проекта. Связи в БД между темами, терминами и фразами организованы в виде тезауруса идей. В каждой теме выделяются главные термины, из которых строится иерархия тем на основе ассоциативных связей главных терминов. Также определяются ключевые фразы для каждой темы.

Организация коллекции текстовых документов в виде кластеров производится так, чтобы документы в пределах каждого кластера были похожи. Текст обычно отображается в векторном пространстве (в виде «мешка слов»), и каждый документ представляется как вектор признаков с использованием схемы взвешивания. Затем выполняется кластеризация путем измерения расстояния между векторами признаков. Тематические модели основаны на идее, что документы представляют собой смеси тем, где тема является распределением вероятностей по словам [57]. Список ключевых слов темы может быть получен таким путем, но извлекаемые темы надо сопровождать процедурой кластеризации, так как тематические модели, в принципе, не предназначены для кластеризации [58].

При автоматическом анализе текстов, предположительно имеющих экстремистскую направленность, существует проблема идентификации её наличия. Предполагается, что система справляется с анализом большинства текстов путем лексического анализа и может выделять относительно редкие случаи, где её «интеллекта» недостаточно для надежного распознавания. Эти случаи

передаются системой сертифицированным экспертам для дальнейшего анализа.

3.3. Расчёт меры семантического подобия текстов. Тематическое моделирование

Авторами предложен показатель *меры подобия двух произвольных текстов*. Для настройки её параметров используется корреляция с наличием формальной ссылки между текстами [5]. Оптимальные параметры алгоритма расчета находятся в ПО, имеющей много формальных ссылок; затем они используются в других ПО в соответствии с гипотезой о близости оптимальных параметров алгоритма расчета меры подобия и неявных ссылок в различных ПО.

Мера подобия — это отношение суммы весов общих значимых словосочетаний (ЗС) и их семантических эквивалентов к сумме весов ЗС и их эквивалентов в меньшем тексте. Мера подобия текста самого на себя равна 100%. В простейшем случае вес ЗС принимается равным количеству символов в ЗС. Чем больше мера подобия текстов, тем выше вероятность наличия явной ссылки между ними. Зависимость была измерена экспериментально и представлена на следующем графике зависимости меры связанности от меры подобия — нелинейная шкала подобия с равномерным размещением пар (рис. 1). При семантическом подобии >8% вероятность формальной связанности >80%.

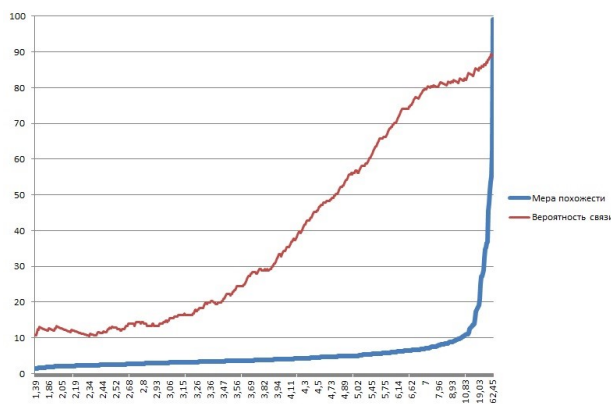


Рис. 1: Зависимость вероятности формальной связанности статей от их семантического подобия

Мера подобия (мера неявной связанности) выражает наличие двух общих терминов, фраз и ЗС. Она рассчитывается по формуле:

$$M = \frac{S_{intersect}}{\min(S_1, S_2)} \cdot 100\%,$$

где: $S_{intersect}$ — сумма весов всех значимых словосочетаний и их семантических эквивалентов, входящих в оба текста, S_1 — сумма весов ЗС и их эквивалентов первого текста, а S_2 — второго.

Для построения графика взяты 119 напрямую связанных пар текстов (мера связанности 100%), 112 косвенно связанных пар (мера связанности 80%) и 350 несвязанных пар (мера связанности 0). Точки на горизонтальной оси соответствуют парам текстов (581 пара) и значения на горизонтальной оси отражают меры подобия этих пар. Мера подобия выбранных пар меняется от 1,39% до 99%. По вертикали для каждой пары отложена мера подобия (синяя линия) и вероятность наличия связи в паре (красная линия).

Большинство пар имеют меру подобия менее 4% и вероятность наличия связи менее 30%. Можно считать, что между двумя текстами есть неявная ссылка/связь, если их мера подобия составляет более 8%. В этом случае вероятность наличия связи по явным ссылкам между этими текстами более 80%. Мера подобия помогает находить пропущенные связи между текстами.

Для улучшения вычисления ИИВ требуется совершенствование алгоритма выявления неявных ссылок с использованием корреляции формальных и неявных ссылок в качестве критерия качества алгоритма. Тексты (Сайты) считаются связанными, если между ними есть прямая ссылка (мера связанности равна 100%) или если они связаны через другой текст (мера связанности равна 80%). Данные величины меры связанности выбраны при максимизации корреляции. На рис. 2 представлен график зависимости корреляции от мини-

мальной длины ЗС, которые рассматриваются при расчете корреляции.

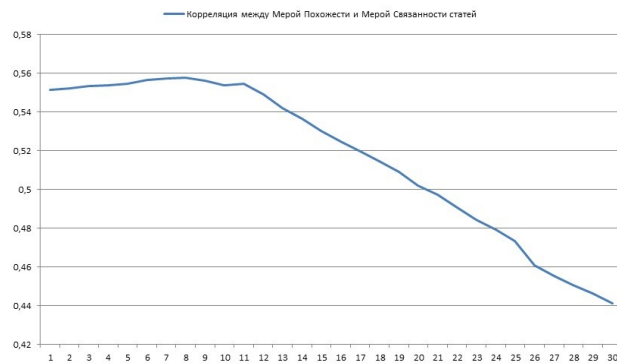


Рис. 2: Максимизация корреляции в зависимости от порога длины фрагментов текста

График показывает, что максимум корреляции 55,7% достигается, если при расчете меры подобия игнорируются словосочетания с длиной меньше 8 букв. Игнорирование более длинных словосочетаний приводит к уменьшению корреляции. Корреляция рассчитана при игнорировании 5% наиболее похожих пар статей. Эти пары должны быть связаны между собой формально, т. к. содержат слишком много общих фраз и идей. Значение «8» параметра минимальной длины ЗС найдено оптимальным для ПО «компьютерная графика» [8] и в соответствии с гипотезой о близости оптимальных параметров может быть использовано в других предметных областях.

Для автоматического выявления сходных фраз и ЗС используется авторский лингвистический процессор; методы тематического анализа (LDA, LSA, ARTM); а также метод построения ассоциативного портрета предметной области (АППО), позволяющий путем анализа больших объёмов текстовых данных (Big Data) автоматически выявлять значимые термины и ассоциативные связи между ними. Идеи (концепты) могут быть адекватно выражены множеством сходных фраз/терминов; близким аналогом является один из методов ТМ: метод латентного размещения Дирихле — LDA. Модель Дирихле — это порождающая модель, рассматривающая каждый до-

кумент как смесь небольшого количества тем и связывающая появление каждого слова с одной из тем документа. Она используется для обработки неструктурированных текстов, позволяя автоматически классифицировать новые тексты, оценивать сходство между ними и создавать вероятностную модель большой текстовой коллекции.

В создаваемой авторами БД с помощью программы-классификатора текстов на основе ассоциативных связей и меры подобия решается задача группирования предметных областей и ключевых слов по категориям. Перед выделением тем текстовая коллекция подвергается предобработке, выделяющей только тематически значимые слова (ЗС) в каждом документе. Для улучшения параметров иерархической кластеризации ЗС используются модели LDA и ARTM совместно с моделями дистрибутивной семантики и семантических векторных пространств. Смысловое сходство фраз определяется с помощью грамматических трансформаций, переводных эквивалентов, синонимии и функциональной синонимии, а также с помощью ассоциативных связей, выявленных по авторской методике АППО [15, 21, 22, 56], которая в данном проекте используется для изучения статистических свойств коллекции текстов. Для расчета ассоциативной связи (семантической близости) между ЗС/терминами ПО используется показатель «косинусная мера», сравнивающая параметры контекста ЗС (контекстные вектора). Иерархические связи между ЗС выбираются из числа ассоциативных связей таким образом, чтобы более общие ЗС имели большее количество ассоциативных связей, при этом учитываются лексико-синтаксические шаблоны, объединяющие данные ЗС. Для вычисления косинусной меры применяется следующая формула:

$$\frac{x \cdot y}{|x||y|} = \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}},$$

где: x и y — вектора в пространстве признаков, которыми являются сами значимые

словосочетания (ЗС) из данного семантического контекстного пространства, а i — это индекс, пробегающий количество признаков.

Для структуризации текстов и построения иерархии категорий применяется методика построения Ассоциативно-Иерархического Портрета Предметной Области (АИППО), где для расчета иерархических связей между ЗС/ЗТ используются методы тематического моделирования, такие, как LDA, hLDA и ARTM [15]. Выделенные по указанным методикам ассоциативные и иерархические связи между ЗС позволяют разрабатывать более совершенные методы и меры подобия исследуемых текстов.

Методы тематического моделирования (ТМ) служат для построения тематической модели коллекции документов. Тематическая модель описывает каждую тему дискретным распределением на множестве терминов, а каждый документ — дискретным распределением на множестве тем. Задача построения тематической модели ставится следующим образом: задана коллекция текстовых документов D . Каждый документ d из коллекции D представляет собой последовательность терминов $W_d = (w_1, \dots, w_{nd})$ из словаря W , где nd — длина документа d . Предполагается, что каждый документ может относиться к одной или нескольким темам. Темы отличаются друг от друга различной частотой употребления слов.

Предполагается, что существует конечное множество тем T , и каждое употребление термина w в каждом документе d связано с некоторой темой $t \in T$, которая не известна. Коллекция документов рассматривается как множество троек (d, w, t) , выбранных случайно и независимо из дискретного распределения $p(d, w, t)$, заданного на конечном множестве $D \times W \times T$. Документы $d \in D$ и термины $w \in W$ являются наблюдаемыми переменными, тема $t \in T$ является латентной (скрытой) переменной. Построить тематическую модель коллекции документов D — означает найти множество тем T , распределения $p(w|t)$ для всех тем $t \in T$ и

распределения $p(t|d)$ для всех документов $d \in D$. Также принимается гипотеза условной независимости, говорящая о том, что появление слов в документе d , относящихся к теме t , описывается общим для всей коллекции распределением $p(w|t)$ и не зависит от документа d , то есть $p(w|d, t) = p(w|t)$.

3.4. Лингвистический процессор неявных ссылок (ЛПНС). Метод релевантных фраз

Суть метода релевантных фраз состоит в поиске наиболее информативных фрагментов в текстах ПО, которая задаётся множеством ключевых терминов с дальнейшим переходом к релевантным этому множеству текстам и формированию ассоциативного портрета ПО. Для оценки информативности фразы подсчитывается индекс встречаемости этой фразы и её семантических эквивалентов в других документах. Рейтинг/значимость Сайта вычисляется с учетом входящих в него информативных фраз, количества формальных ссылок (гиперссылок) и неявных контекстных ссылок.

Для анализа формальных ссылок в статьях коллекции (в БД) используется авторский лингвистический процессор BREF (Book Reference) [6], который является частью Лингвистического Процессора Неявных Ссылок (ЛПНС, версия 1.1). ЛПНС автоматически находит в текстах формальные связи и представляет их в формате простого самодокументируемого XML-файла (для настройки оптимальных параметров меры семантического подобия и индекса идеологического влияния ИИВ). ЛПНС позволяет по выделенной информации построить Граф цитирования и Граф соавторов, отражающих формальные связи в БД. ЛПНС настраивается по корпусу текстов путем максимизации корреляции между формальными и неявными ссылками.

С помощью ЛПНС был проведён анализ текстов коллекции БД, по результатам которого сделаны следующие выводы: 1) неявные связи существуют и могут быть выделены из текстов путем сравнения их

лексики с лексикой посторонних коллекций; 2) явная ссылка почти всегда сопровождается группой неявных ссылок; 3) может быть построен рейтинг слов и словосочетаний/фраз, потенциально являющихся неявными ссылками, в котором наиболее вероятные кандидаты будут занимать лидирующие позиции.

Были предложены формальные правила отбора ЗС из рейтинга словосочетаний, в частности, в виде порога по частоте во внешней коллекции. Весовая функция подбиралась по максимуму корреляции между явными и неявными ссылками, при этом наиболее значимые словосочетания получили наибольший вес и попали в начало списка.

3.5. Расчёт индекса ИИВ. Анализ неявных ссылок

Данная методика позволяет ранжировать Сайты по степени влияния (*impact*) и служит основой для создания программы расчета индекса идеологического влияния (ИИВ), а также основанного на ИИВ прогнозного анализа формирования и роста экстремистских групп.

ИИВ рассчитывается по количеству неявных связей (ссылок) между Сайтами, которые представляются в виде множества значимых словосочетаний (ЗС). Неявная ссылка может быть представлена одним словом/фразой/ЗС или набором близких по смыслу слов/фраз/ЗС. При выделении неявных ссылок используются: грамматические трансформации; преобразование фраз на основе синонимии; изменение словарных форм; сопоставление с онтологиями; программы перевода; терминологический анализ; окрестностный подход (учет ассоциативных и иерархических связей). Расчет ИИВ производится с помощью введенного авторами показателя — меры подобия связанных между собой текстов (Сайтов) с учетом времени и некоторых других факторов/параметров (см. рис. 2,3). Оптимальные параметры алгоритма выявления неявных ссылок подбираются на те-

матической коллекции с большим количеством формальных ссылок между документами (эталонном служила коллекция по компьютерной графике) путем максимизации корреляции явных и неявных (ИИВ) связей. Затем найденные значения параметров используются для выделения неявных ссылок в создаваемой БД текстов экстремистской тематики для анализа выбранных тематических подкатегорий (например, «кафиры» и «ватники»). Основные параметры при исследовании корреляции — это длина фраз/ЗС, количество их вхождений в статью, в тематический корпус и во внешний посторонний корпус.

Рассмотрим ИИВ подробнее. В 2016 г. авторами был предложен и апробирован алгоритм автоматического выявления неявных ссылок, на основании которого рассчитывается Индекс Контекстного Научного Цитирования (ИКНЦ), позволяющий проводить оценку качества новых статей в различных ПО [3, 10]. Алгоритм использует введенную авторами меру семантического подобия текстов, которая настраивается по максимуму корреляции между прямыми и неявными ссылками. В ходе экспериментов (объем коллекции 900 научных документов) на ПО «Компьютерная графика, визуализация и виртуальное окружение в СССР и России (КГВ)» был найден один из параметров алгоритма (минимальная длина анализируемых слов), при котором корреляция достигает максимума для статей в ПО КГВ.

Принципы построения ИКНЦ были положены в основу разработки индекса идеологического влияния (ИИВ). Алгоритм расчета индекса ИИВ Сайтов работает сходно с алгоритмом ИКНЦ: величина ИИВ зависит от количества неявных связей и строится на основе лингвистического процессора (ЛПНС), выделяющего неявные связи между Сайтами (текстами) с помощью меры семантического подобия. Апробация методики на конкретных примерах — анализ коллекций на тему «ватники-москалы», «ка-

фигуры» — подробно изложены в статьях [2, 5].

3.6. Методы визуального анализа топологической структуры взаимосвязей экстремистских группировок

Визуальные представления на основе полученных из Интернет данных часто строятся на основе методов статистики и многомерного масштабирования (multidimensional scaling), выполненных с помощью R-скриптов (<http://cran.r-project.org>). Однако существует более 300 способов визуализации больших текстовых коллекций, не поддающиеся единой классификации (формат входных данных, заданные параметры и т. д.). Примеры практического применения авторами визуализации тематических связей — анализ собранных текстовых коллекций, в том числе с помощью системы IQBuzz, тематическое моделирование LDA с помощью пакета BigARTM и последующая визуализация программой Gephi [59], а также анализ полученных гистограмм и графов детально рассмотрены в статьях [7, 12, 13, 22]. В результате были обнаружены тревожные события (рост употребления терминов-антагонистов), предшествующие росту экстремистской активности; построены графики динамики употребления терминов, например, выявлено начало роста использования термина «кафир» в английском языке в 1970-х годах (рис. 3,4).



Рис. 3: Количество записей в социальных сетях с термином-антагонистом «Кафир» в день (с 2014-01-01 по 2017-03-01)

Виден рост в 2015 и 2016 годах. Спад в 2017 объясняется неполными данными

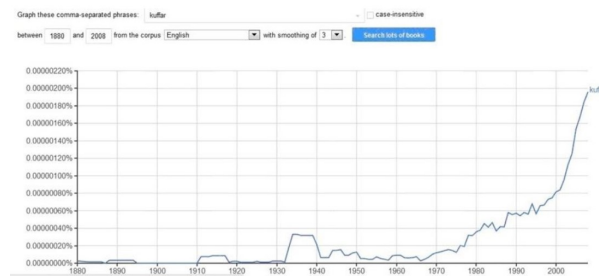


Рис. 4: Динамика роста употребления термина-антагониста “kuffar” в книгах на английском языке. Этот рост — **тревожное событие**, по которому уже в 1977 году (40 лет назад) можно было прогнозировать рост террористической активности и вероятность будущих терактов (ИГИЛ)

4. Заключение

Авторским коллективом получены следующие важнейшие результаты: 1) разработаны теоретические и практические основы комплексного подхода (совокупность методов ТМ, вероятностно-статистического, дистрибутивного анализа, корпусной лингвистики, визуального анализа и др.), применяемого на текстовых документах экстремистской тематики; 2) разработана концепция автоматизированной технологии для мониторинга, прогноза и визуализации экстремистской активности в Интернет; определён состав её основных модулей; 3) разработаны программные средства (первая итерация), обеспечивающие работу модулей; 4) создан и совершенствуется в ходе машинного обучения алгоритм построения и постоянного пополнения БД текстов из Интернет по экстремистской тематике; 5) построены коллекции открытых документов/текстов из Интернета объемом более 10 Гб; 6) составлены коллекции размером более 10000 датированных документов/текстов из соцсетей и блогов ПО «Терроризм»; 7) разработана первая версия лингвистического процессора (ЛПНС) для выявления неявных ссылок с заданными параметрами; 8) предложена и апробирована новая мера семантического подобия текстов, с помощью которой определяется наличие неявной ссылки между текстами; 9) предложен и частично апробирован новый алгоритм расчета обладающего прогнозными свойствами индекса идеологического влияния (ИИВ) Сайтов; 10) предложена и ча-

стично апробирована гипотеза о близости оптимальных параметров неявных ссылок в различных текстовых корпусах и предметных областях.

Значимость достигнутого заключается в создании автоматизированного инструмента, способного не только выявлять существующие сайты/группы экстремистской направленности и отдельные лица (аккаунты) в соцсетях, но и прогнозировать формирование подобных сообществ в Интернете.

Многоязычный индекс ИИВ позволяет находить значимые документы и источники происхождения идей независимо от языка. Разрабатываемая методика позволяет выявлять тексты–первоисточники различных идей в социальных сетях и в Интернете, позволяет проследивать потоки распространения идей, а также быстрее находить достоверную и качественную информацию из первоисточников. Межъязыковое определение семантического текстового подобия является важным шагом для обнаружения и оценки межъязыкового плагиата, исследования в данном направлении редки.

Литература

- [1] Золотарев О.В., Шарнин М.М., Клименко С.В. Семантический подход к анализу террористической активности в сети Интернет на основе методов тематического моделирования. // Вестник Российского нового университета. Серия: Сложные системы: модели, анализ и управление. М.: РосНОУ, 2016. №3. С. 64-71.
- [2] Клименко С.В., Шарнин М.М., Ищенко Н.С., Хакимова А.Х. Использование метода неявных ссылок для противодействия распространению деструктивных антигуманистических идеологий в сети Интернет // Международная Школа-семинар «Ситуационные центры и информационно-аналитические системы класса 4i для задач мониторинга и безопасности» SCVRT2015-16, Пущино, ЦарьГрад, 21-24 ноября 2015-2016 гг., С. 129-133.
- [3] Клименко С.В., Шарнин М.М., Хакимова А.Х., Золотарев О.В., Мацкевич А.Г. Методы оценки качества и влияния (iprast) научных статей для повышения объективности индекса научного цитирования. // Вестник Российского нового университета. Серия: Сложные системы: модели, анализ и управление. М.: РосНОУ, 2016. №3. С. 51-59.
- [4] Шарнин М.М., Сомин Н.В. Об основных типах связанности между текстовыми документами. // Системы и средства информатики, 2017. Т. 27, №1, С. 104-111.
- [5] Хакимова А.Х., Шарнин М.М., Клименко С.В., Золотарев О.В., Родина И.В. Мера подобия текстов как инструмент оценки интертекстуальности при анализе больших коллекций документов. // Вестник Российского нового университета, М.: РосНОУ, 2016. — Выпуск 4, С. 62-71.
- [6] Золотарев О.В., Шарнин М.М., Клименко С.В., Кузнецов К.И. Система PullEnti — извлечение информации из текстов естественного языка и автоматизированное построение информационных систем // Сборник трудов Международной конференции «Ситуационные центры и информационно-аналитические системы класса 4i для задач мониторинга и безопасности» SCVRT2015-16, Пущино, ЦарьГрад, 21-24 ноября 2015-2016, Пущино, ЦарьГрад, 21-24 ноября 2015-2016 гг., С. 28-35.
- [7] Шарнин М.М., Клименко С.В., Слепцов И.О. Технология визуализации графа соавторов по открытой информации из Интернета // Сборник трудов Международной конференции «Ситуационные центры и информационно-аналитические системы класса 4i для задач мониторинга и безопасности» SCVRT2015-16, Пущино, ЦарьГрад, 21-24 ноября 2015-2016 гг., С. 168-174.
- [8] Орлов Д.Е., Клименко С.В., Орлова Н.А., Шарнин М.М., Слепцов И.О. Построение и анализ тематических коллекций научных публикаций по тематике «Компьютерная графика, визуализация и виртуальная реальность» // Сборник трудов Международной конференции «Ситуационные центры и информационно-аналитические системы класса 4i для задач мониторинга и безопасности» SCVRT2015-16, Пущино, ЦарьГрад, 21-24 ноября 2015-2016 гг., С. 36-42.
- [9] Слепцов И.О., Косиков С.В. Потенциально-бесконечные списки: средства для обработки упорядоченных структур. Практическое применение. // Труды Международной научной Школы-семинара «Ситуационные центры и ИАС4i для задач мониторинга и безопасности» SCVRT1516 21-24 ноября 2015-2016, ЦарьГрад, Московская обл., Россия ISBN 978-5-88835-043-0. Институт физико-технической информатики (ИФТИ). 2016.
- [10] Клименко С.В., Шарнин М.М., Орлов Д.Е., Орлова Н.А., Мацкевич А.Г. Исследование корреляции библиографических и неявных ссылок для определения значимых документов и перспективных направлений развития. // Труды Международной научной Школы-семинара «Ситуационные центры и ИАС4i для задач мониторинга и безопасности» SCVRT2015-16 21-24 ноября 2015-2016, ЦарьГрад, Московская обл., Россия ISBN 978-5-88835-043-0. Институт физико-технической информатики (ИФТИ). 2016. С. 1-5.
- [11] Потапова З.Е., Протасов В.И., Шарнин М.М. Использование коллективного интеллекта экспертов в сложных случаях анализа текстов, предположительно имеющих экстремистскую направленность.

- // Сборник трудов Международной конференции СРТ2017, 07-14 мая 2017, Ларнака, Кипр. Изд-во Института физико-технической информатики (ИФТИ). 2017. С. 372-374.
- [12] Меркурьева Н.А., Шарнин М.М., Слободюк Е.А., Цыганов В.В. Построение визуальных карт предметной области для разведочного поиска (exploratory search). // Сборник трудов Международной конференции СРТ2017, 07-14 мая 2017, Ларнака, Кипр. Изд-во Института физико-технической информатики (ИФТИ). 2017. С. 271-274.
- [13] Цыганов В.В., Шарнин М.М., Родина И.В. Построение визуальных карт предметной области с помощью методов тематического моделирования по данным социальных сетей. // Сборник трудов Международной конференции СРТ2017, 07-14 мая 2017, Ларнака, Кипр, Изд-во Института физико-технической информатики (ИФТИ). 2017. С. 355-359.
- [14] Демидов А.О., Шарнин М.М. Построение развивающейся во времени системы тематических категорий для пополняемого корпуса текстов. // Сборник трудов Международной конференции СРТ2017, 07-14 мая 2017, Ларнака, Кипр. Изд-во Института физико-технической информатики (ИФТИ). 2017. С. 343-349.
- [15] Шарнин М.М., Клименко С.В., Хакимова А.Х., Сюракшина Ю.В. Методика построения ассоциативно-иерархического портрета предметной области. // Сборник трудов Международной конференции СРТ1617, 07-14 мая 2017, Ларнака, Кипр. Изд-во Института физико-технической информатики (ИФТИ). 2017. С. 343-349.
- [16] Цимбалов А.В., Золотарев О.В. Метод шинглов. // Вестник Российского нового университета, М.: РосНОУ, 2016. — Выпуск 4. С. 72-79.
- [17] Золотарев О.В., Шарнин М.М., Клименко С.В. Использование онтологического подхода для анализа текстов естественного языка. // Вестник Российского нового университета. Серия: Сложные системы: модели, анализ и управление. 2017. №1. (в печати).
- [18] Michael Charnine, Stanislav Klimenko. Semantic cyberspace of scientific papers. // 2017 International Conference on CYBERWORLDS. University of Chester, Chester, United Kingdom. 20-22nd September, 2017. In cooperation with Eurographics Association, ACM SIGGRAPH, and the International Federation for Information Processing (IFIP) — <http://www.cyberworlds-conference.org/> (в печати).
- [19] Charnine M., Somin N. and Nikolaev V. Conceptual Text Generation Based on Key Phrases. // Proceedings of the 2014 International Conference on Artificial Intelligence (ICAI 2014). Vol.I. WORLDCOMP'14. July 21-24, 2014. Las Vegas Nevada, USA. CSREA Press. С. 639-643.
- [20] Charnine M., Klimenko S. Measuring of “Idea-based” Influence of Scientific Papers // Proceedings of the 2015 International Conference on Information Science and Security (ICISS 2015), December 14-16, 2015. — Seoul, South Korea. — Pp. 160-164.
- [21] Galina I., Charnine M., Somin N., Nikolaev V., Morozova Yu., Zolotarev O. “Metod for generating subject area associative portraits: different examples”. // Proceedings of the 2015 International Conference on Artificial Intelligence (ICAI 2015), WORLDCOMP'15, July 27-30, 2015. Las Vegas Nevada, USA, v.I, pp. 288-294.
- [22] Галина И.В., Козеренко Е.Б., Морозова Ю.И., Сомин Н.В., Шарнин М.М. Ассоциативные портреты предметной области (аппо) — инструмент автоматизированного построения систем big data для извлечения знаний: теория, методика, визуализация аппо, возможное применение. // Информатика и её применения. 2015. Т. 9. Вып. 2. С. 94-111.
- [23] Emilio Ferrara, Wen-Qiang Wang, Onur Valor, Alessandro Flammini, Aram Galstyan (2016) Predicting online extremism, content adopters, and interaction reciprocity // arXiv:1605.00659 [cs.SI].
- [24] Johnson N.F., Zheng. M., Vorobyeva Y., Gabriel A., Qi H., Manrique P., Johnson D., Restrepo E., Song C., Wuchty S. New online ecology of adversarial aggregates: Isis and beyond // Science, 17 Jun 2016, pp. 1459-1463.
- [25] Törnberg A., Törnberg P. Muslims in social media discourse: Combining topic modeling and critical discourse analysis // Discourse, Context and Media 13, 2016, PP. 401-406.
- [26] Vighne S., Trimbake P., Musmade A., Merukar A., Pandit S. An Approach to Detect Terror Related Activities on Net // IJARIE, Vol-2, Issue-1, 2016. PP. 401-406.
- [27] Towne, W. Ben, Rosé, Carolyn P., and Herbsleb, James D. Measuring Similarity Similarity: LDA and Human Perception // ACM Trans. Intell. Syst. Technol. 7, 2, Article 25, 2016, pp. 1-29.
- [28] Wang S. and Koopman R. Semantic embedding for information retrieval // BIR 2017 Workshop on Bibliometric-enhanced Information Retrieval; 2017, pp. 122-131.
- [29] Tomas Mikolov, Greg Corrado, Kai Chen, and Jeffrey Dean. Efficient Estimation of Word Representations in Vector Space // International Conference on Learning Representations (ICLR 2013), 1 2013, pp. 1-12.
- [30] Jey Han Lau, Timothy Baldwin. An Empirical Evaluation of doc2vec with Practical Insights into Document Embedding Generation // 1st Workshop on Representation Learning for NLP, Berlin, Germany, 2016, pp. 78-86.
- [31] Kim S., Wilbur W.J., Lu Z. Bridging the Gap: a Semantic Similarity Measure between Queries and Documents // In ArXiv e-prints, 2016.
- [32] Abdelwahab O., Elmaghraby A. UoFL at SemEval-2016 Task4: Multi Domain word2vec for Twitter Sentiment Classification // SemEval-2016, pp. 169-175.

- [33] Xu S., Liang H. and Baldwin T. UNIMELB at SemEval-2016 Tasks 4A and 4B: An Ensemble of Neural Networks and a Word2Vec Based Model for Sentiment Classification // SemEval-2016, pp. 168-194.
- [34] Junfeng Tian, Man Lan. ECNU at SemEval-2016 Task 1: Leveraging Word Embedding from Macro and Micro Views to Boost Performance for Semantic Textual Similarity // SemEval-2016, pp. 618-624.
- [35] Середнев В.А. Интернет как эффективное сетевое пространство для проведения оперативно-розыскных мероприятий. // Вопросы современной юриспруденции: сб. ст. по матер. LX междунар. науч.-практ. конф. №4(55). — Новосибирск: СибАК, 2016. — С. 67-80.
- [36] Воронцов С.А. Терроризм 2016: новые вызовы и угрозы. // Журнал Северо-Кавказский юридический вестник. Выпуск №3. 2016. С. 108-113.
- [37] Чурилов С.А., Быкадорова А.С. О противодействии распространению и профилактике радикальных идеологий в молодежной среде и сети Интернет. // Обзор нцпги. Выпуск №8. 2016. С. 39-43.
- [38] Горбачев А.М. Big data как инструмент противодействия угрозам экстремизма Международно-правовые средства противодействия терроризму в условиях глобализации. Проблемы террористического наемничества среди молодежи и пути их преодоления // Сб. материалов Всероссийской научно-практ. конф. — Ставрополь: СГПИ, 2016
- [39] Рязанов Д.С., Охрименко А.В., Ибрагимов Г.И. Терроризм в странах мусульманского культурно-исторического ареала: общие тренды и идеологические особенности (по материалам анализа электронных баз данных терроризма) // Исламоведение. 2016. Том 7. №1 (27). <http://islam.dgu.ru>
- [40] Кирченко Е.Г. Правовое регулирование и противодействие экстремизму в сети Интернет // Противодействие процессу вовлечения молодежи в деятельность экстремистских и террористических организаций // Материалы конференции 27 апреля 2016 года, Брянск: РИО БГУ, 2016, С. 10-15.
- [41] Басараб М.А., Иванов И.П., Колесников А.В., Матвеев В.А. Обнаружение противоправной деятельности в киберпространстве на основе анализа социальных сетей: алгоритмы, методы и средства (обзор) // Вопросы кибербезопасности, №4(17), 2016. С. 11-20.
- [42] Суслов С.И. Сетевые агенты политической интернет коммуникации в русскоязычном онлайн пространстве. Диссертация к.с.н., С-Пб, 2017, 290 С.
- [43] Анянзева М.И., Девяткин Д.А., Кобозева М.В., Смирнов И.В. Лингвостатистический анализ текстов экстремистской направленности // Тр. Междунар. науч. конф. Московского физико-технического института (государственного университета) и Института физико-технической информатики (SCVRT1516). М.: Протвино, 2016. С. 210-213.
- [44] Анянзева М.И., Кобозева М.В., Соловьев Ф.Н., Поляков И.В., Чеповский А.М. О проблеме выявления экстремистской направленности в текстах // Вестн. Новосиб. гос. ун-та. Серия: Информационные технологии. 2016. Т. 14, №4. С. 5-13.
- [45] Воронцов К.В. Лекции по тематическому моделированию. 2013г. <http://www.machinelearning.ru/wiki/images/2/22/Voron-2013-ptm.pdf>
- [46] Воронцов К.В. тематическое моделирование. 2016 г. <http://www.machinelearning.ru/wiki/images/e/e6/Voron-ML-TopicModeling-slides.pdf>
- [47] Савельев В.В. Применение методов тематического моделирования к профессиональным стандартам для реализации модульного подхода к классификации профессий. // Педагогическое образование в России. 2017. № 3. С. 22-28.
- [48] Янина А.О., Воронцов К.В. Мультимодальные тематические модели для разведочного поиска в коллективном блоге. // Машинное обучение и анализ данных. 2016. Том 2, № 2. С. 1-14.
- [49] Машечкин И.В., Петровский М.И., Поспелова И.И., Царев Д.В. Методы автоматического аннотирования и выделения ключевых слов в задаче обнаружения экстремистской информации в сети Интернет. // Современные информационные технологии и ИТ-образование. 2016. Т. 12. № 1. С. 188-200.
- [50] Машечкин И.В., Петровский М.И., Царев Д.В. Методы машинного обучения для анализа поведения пользователей при работе с текстовыми данными в задачах информационной безопасности. // Вестник Московского университета. Серия 15: Вычислительная математика и кибернетика, издательство Изд-во Моск. ун-та (М.), № 4, 2016. С. 33-39.
- [51] Bettencourt L. et al. The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models // Physica A. — 2006. — May. — Pp. 513-536.
- [52] Jacobs J.A. Receptivity Curves: Educational Research and the Flow of Ideas: Expanded Version, Population Studies Center, University of Pennsylvania, PSC Working Paper Series, PSC 13-10. — 2013. — http://repository.upenn.edu/psc_working_papers/50
- [53] Jee J., Klippel L., Hossain M., Ramakrishnan Ramakrishnan N., Mishra B. Discovering the Ebb and Flow of Ideas from Text Corpora // IEEE Computer 45(2). — 2012. — Pp. 73-77.
- [54] IQBuzz.pro — Сервис мониторинга социальных медиа и онлайн СМИ.// <http://iqbuzz.pro/about.php>
- [55] Charnine Michael, "Keywen: Automated Writing Tools", Booktango, USA, 2013, ISBN 978-1-46892-205-9.
- [56] Шарнин М.М., Сомин Н.В., Кузнецов И.П., Морозова Ю.И., Галина И.В., Козеренко Е.Б. Статистические механизмы формирования ассоциативных портретов предметных областей на основе естественно-языковых текстов больших объемов для систем извлечения знаний. // Информатика и её применения. 2013. Т. 7. № 2. Стр. 92-99.

- [57] Steyvers M., Griffiths T. Probabilistic Topic Models. — Hillsdale, NJ, USA : Erlbaum, 2007.
- [58] Kim S., Wilbur W.J. Thematic clustering of text documents using an EM-based approach // Biomed Semantics. — 2012; 3(Suppl 3). — S. 6.
- [59] Mathieu Bastian, Sebastien Heymann and Mathieu Jacomy Gephi: An open source software for exploring and manipulating networks, 2009.

Использование методов тематического моделирования в мультязыковых средах для прогноза тревожных событий*

М. М. Шарнин¹, Н. С. Ищенко², Н. Ю. Пахмутова³, Ю. В. Сюракшина⁴

mc@keywen.com, niofterna@gmail.com, vidsilvestre@gmail.com,
yuliya.surakshina@phystech.edu

Москва, ¹Институт Проблем Информатики ФИЦ ИУ РАН

Луганск, ²Луганская академия культуры и искусств имени М. Матусовского

³Московский государственный университет имени Ломоносова

⁴Московский физико-технический институт (государственный университет)

В статье изложены результаты практического применения методов тематического моделирования в мультязыковых средах для мониторинга экстремистской активности в Интернете и прогноза тревожных событий. При работе с двумя корпусами текстов, содержащих экстремистскую идеологию радикальных мусульман и украинских националистов подбираются оптимальные параметры для метода неявных ссылок, рассчитывается мера подобия корпусов текста, определяется общая и специфическая характерная терминология двух корпусов текстов. Выделение нескольких категорий терминов по степени ожесточенности позволяет определить этапы роста и спада интереса к теме, прогнозировать будущие этапы, и тем самым — тревожные события в реальности.

Ключевые слова: тематическое моделирование, автоматизированные системы, Big Data, индекс идеологического влияния, экстремизм, Интернет, неявные ссылки, метод неявного контекстного цитирования, индекс идеологического влияния, общие характерные термины, специфические характерные термины, прогноз тревожных событий, мультязыковые среды

Topic modeling methods in multi-language environments for troubling events prediction*

M. M. Charnine¹, N. S. Ishchenko², N. Y. Pakhmutova³, Y. V. Surakshina⁴

Moscow, ¹FRC CSC RAS

Lugansk, ² Lugansk State Academy of Culture and Arts named by M. Matusovsky

³Lomonosov Moscow State University

⁴Moscow Institute of Physics and Technology (State University)

The paper exposes the results of practical application of the topic modeling methods in multi-language environment in extremist activity monitoring on the Internet, and in forecasting of troubling events. The study was based on two texts corpora containing extremist ideology of radical Muslims and Ukrainian nationalists respectively. The optimal parameters' selection was performed for the method of implicit references; a measure of the similarity of the text corpora was calculated; common terminology of the two corpora and specific terminology for each one were determined. The selection of several categories of terms according to the degree of their aggressivity allows determining the stages of growth/decline of the attention for a specific topic; it allows predicting the nature of the further stages and thus — the occurrence of the disturbing events in reality.

Keywords: topic modeling, automated systems, Big Data, index of ideological impact, extremism, Internet, implicit references, implicit contextual citing method, general characteristic terms, specific characteristic terms, forecast of troubling events, multi-language environment

1. Введение. Постановка задачи

Прогнозирование тревожных событий на основе данных сетевой активности — сложная задача. Тревожные события включают в себя нагнетание конфликтов в разных сферах общественной жизни, эскалацию насилия, теракты. Сетевая активность,

которая сопровождает появление тревожных событий, направлена на распространение экстремистских идеологий, легитимирующих применение насилия в отношении антагонистов этой идеологии. Авторский коллектив исходит из гипотезы, что тревожные события, вызванные деятельностью определенной экстремистской группы, коррелируют с сетевой активностью этой

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 16-07-00756, 16-29-09527, 18-07-20030, 18-37-10008.

группы или других групп, распространяющих ту же экстремистскую идеологию.

Из-за чрезвычайно широкого толкования термина «экстремизм» в рамках экспертно-юридической практики последних лет, авторы полагают целесообразным обозначить идейно-политическую основу определения этого термина в рамках настоящей статьи и других документов, созданных авторским коллективом.

Под экстремизмом подразумевается манипуляция сознанием аудитории, направленная на оправдание и распространение дискриминационных практик разного рода, оправдание, пропаганду и героизацию прямого насилия по отношению к гражданскому населению РФ и зарубежным соотечественникам (с маркированием их как коллективного антагониста). Более подробно характерные признаки подобных манипуляций и экстремистских групп в сети Интернет описаны в разделах 4.1 и 4.2.

Таким образом, первый шаг на пути прогноза тревожных событий — это выявление групп, распространяющих в сети экстремистские идеологии, оправдывающие насилие. Для выявления таких групп в мультязыковых средах авторский коллектив предлагает использовать методы тематического моделирования (ТМ), оптимальные параметры которых определяются при работе с реально существующими экстремистскими группами, а именно мусульманскими радикалами и украинскими националистами. Обе эти группы имеют экстремистские идеологии, оправдывающие насилие в отношении инакомыслящих, обе эти идеологии широко представлены в сети, обе группы известны реальной экстремистской деятельностью, которая легко отслеживается. Закономерности, найденные на этом материале, уточняют друг друга и позволяют найти общие характеристики сетевой активности экстремистских групп, которые будут использоваться при прогнозе тревожных событий.

2. Краткий обзор современных подходов к проблеме

Исследовать сетевую экстремистскую деятельность в мультязыковых средах довольно сложно. С одной стороны, проявления так называемого «языка вражды» *hate speech* (англ.) [1], которые являются формальным основанием для санкций в рамках собственной цензурной политики Интернет-ресурсов и надзорных органов, зачастую свидетельствуют лишь об эмоциональной несдержанности пользователя в рамках личного конфликта, даже если содержат прямые призывы к насилию. С другой, экстремисты-манипуляторы чаще действуют скрытно, используя форумы с многоуровневым доступом, закрытые аккаунты в социальных сетях и сайты, основная тематика которых не содержит признаков потенциальной пропаганды насилия (например — сетевые сообщества, организованные по гендерному признаку, вроде страниц, объединяющих русскоязычных мусульманок). При этом дискриминация и насилие пропандируются без явного использования языка вражды или при его минимальном использовании.

Другие немаловажные причины сложности исследования — постоянный рост количества текстов и сложность автоматизации тем этой предметной области. Информация, которая используется экстремистскими группами, представлена на разных языках, в разных жанрах, намеренно испорчена или закодирована. Методики поиска по ключевым словам или фразам, связанным с оружием или прямым насилием не работают, поскольку не позволяют отграничить явную и скрытую пропаганду от информационных сообщений.

В зарубежных и отечественных исследованиях для семантического анализа сайтов интересующей нас предметной области используется метод скрытого распределения Дирихле (Latent Dirichlet Allocation, LDA). Наиболее популярные современные методы автоматизации основаны на тематиче-

ском моделировании текстов с использованием латентно-семантического анализа. Латентно-семантический анализ работает с матричным представлением коллекции текстов, получаемым с помощью модели «мешок слов» (англ. *bag-of-words*).

3. Авторская методика — выявление экстремистских сайтов с помощью индекса идеологического влияния

В течение ряда лет авторский коллектив занимается разработкой методов тематического моделирования, которые позволяли бы анализировать сетевую экстремистскую активность. В анализе используются данные, полученные с веб-сайтов, блогов, аккаунтов в социальных сетях. Одним из новых методов ТМ, разрабатываемых нашей группой, является определение индекса идеологического влияния методом неявного контекстного цитирования.

Индекс идеологического влияния — это показатель степени экстремизма сайтов, обладающий прогнозными свойствами. Расчёт ИИВ основан на вероятностной модели влияния (*impact*) тем и идей сайта на темы и идеи сайтов схожей тематики. Поскольку темы и идеи любого текста могут быть адекватно выражены множеством сходных по смыслу лексических фрагментов, можно определить вклад каждого сайта в возникновение и распространение интересующих исследователя тем. Таким образом обнаруживаются скрытые связи между различными сайтами/авторами, определяется источник распространения экстремистской идеологии и возможных тревожных событий.

Авторским коллективом разработана методика нахождения семантического подобия между текстами на разных языках и апробирована релевантность использования ИИВ для анализа экстремистских сайтов в мультязыковых средах.

3.1. Теоретические основы авторской технологии

Базовые элементы авторской автоматизированной технологии подробно рассмот-

рены в других работах авторов данного проекта. Основой метода латентно-семантического анализа, который используется в тематическом моделировании, является анализ явных и неявных ссылок между статьями [2, 3, 4]. Те же подходы применимы и к анализу экстремистских сайтов [5, 6, 7, 8, 9].

Экстремистская идея семантически представляется в виде множества сходных по смыслу фраз или терминов, которые выявляются методами тематического моделирования и дистрибутивной семантики. Извлекая из текстов экстремистские идеи и наблюдая их динамику во времени, можно определить вклад каждого сайта в распространение выявленной темы и тем самым обнаружить скрытые связи между сайтами.

Лексическая база выступает в качестве источника выявления экстремистских сайтов и обнаружения явных и неявных ссылок между ними. На основании выявленных таким образом связей строится индекс идеологического влияния Сайта (ИИВ). Рост идеологического влияния группы в Интернете способствует росту численности этой группы и сопровождается ростом числа сайтов этой группы. ИИВ рассчитывается автоматически по неявным контекстным ссылкам на сайт. Таким образом, использование ИИВ позволяет делить новые сайты на группы, ранжировать их по степени влияния, находить источник экстремистской идеологии и прогнозировать тревожные события.

Анализ лексики на разных языках и за длительный период позволяет обнаружить динамику употребления термина и найти корреляцию с тревожными событиями в разное время. На рис. 1 представлен график употребления термина «кафир» в англоязычных документах.

Динамика употребления термина показывает его рост в некоторые периоды. Этот рост — тревожное событие, по которому уже в 1977 году (40 лет назад) можно было прогнозировать рост террористической ак-

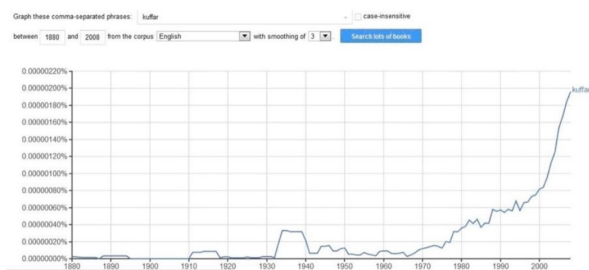


Рис. 1: Динамика роста употребления термина-антагониста "kuffar" в книгах на английском языке

тивности и вероятность будущих терактов (ИГИЛ).

3.2. База данных (БД) и тезаурус идей

Лексические ресурсы проекта включают в себя коллекции текстов, списков тем, словарей терминов, значимых словосочетаний (ЗС) и ключевых фраз. Для хранения лексических ресурсов проекта используется БД, связи в которой организованы в виде тезауруса идей.

Документы представляют собой смеси тем, где тема является распределением вероятностей по словам [10]. В каждой теме выделяются главные термины, из которых строится иерархия тем на основе ассоциативных связей главных терминов. Также определяются ключевые фразы для каждой темы.

В ходе работы текстовые документы были организованы в виде кластеров. Кластером называется файл, содержащий некоторый текст, а также два или более других текстов, которые формально ссылаются на первый. Документы в пределах каждого кластера сгруппированы по схожести. Текст отображен в векторном пространстве в виде «мешка слов», причём каждый документ представлен как вектор признаков с использованием схемы взвешивания. Измерение расстояния между векторами признаков стало основой кластеризации.

3.3. Направленный семантический поиск

Алгоритм поиска и сбора данных включает в себя общий мониторинг Интернета и специализированный мониторинг соцсетей.

Сбор данных из соцсетей осуществляется с помощью системы IQBuzz [11]. Сбор данных из Интернет осуществляется авторской системой «KeyCrawler Интернета для систем извлечения знаний» [12, 13, 14, 15].

Направленный семантический поиск позволяет выявить в сети те сайты, которые распространяют экстремистские идеологии. В работе исследовались документы, содержащие две идеологии — радикальных мусульман и украинских националистов, общим признаком которых является формирование коллективной элитарности. Этот процесс может быть отслежен методами корпусной лингвистики, а именно направленным семантическим поиском по терминам-антагонистам. Когда выделены две коллекции документов, содержащие термины-антагонисты каждой группы, в них определяются разные темы экстремистского характера, и соответственно выделяются общие характерные термины (ОХТ) и специфические характерные термины (СХТ). На следующем этапе характеризуется динамика обеих групп и определяется корреляция динамических закономерностей с реальными тревожными событиями.

4. Обоснование методики работы с группами характерных терминов для прогноза тревожных событий

При исследовании исходим из того, что экстремистский сайт создан экстремистской группой, то есть организацией с единым координационным центром, имеющей один или несколько сайтов/сообществ в сети Интернет и социальных сетях, а также имеющей в реальной жизни связи, не отраженные в Интернет-активности. Экстремистская группа обладает специфическими признаками, и самый важный из них — формирование коллективной элитарности участников.

4.1. Формирование коллективной элитарности как элемент экстремистской идеологии

Коллективная элитарность представляет собой комплекс идей, создающий у носителя чувство избранности по признаку принадлежности к некоторой группе. Формирование коллективной элитарности осуществляется на базе уже сформированной коллективной идентичности и имеет следующие характеристики:

1. Выделение группы в коммуникативном пространстве
2. Появление постоянного самоназвания с положительным смыслом
3. Выделение в коммуникативном пространстве Коллективного Антагониста
4. Появление постоянного названия для Коллективного Антагониста и его дегуманизация

Формирование групп в пространстве коммуникаций происходит постоянно и является естественным процессом. Коллективная идентичность — необходимый признак группы. Первым шагом на пути формирования коллективной идентичности является деление на *мы* и *не-мы*, на принадлежащих к группе и всех, кто за её пределами.

Само по себе разделение людей на две группы и формирование коллективной идентичности по признаку принадлежности к определённому сообществу не является чем-то исключительным, опасным и ведущим к противоправным действиям. Такие сообщества могут как развиваться далее в конструктивном ключе, так и послужить основой для формирования экстремистских групп.

Признаком развития группы в деструктивном направлении является присутствие в коллективной идентичности группы Коллективного Антагониста. Коллективный Антагонист — сообщество, представители которого являются носителями антиценностей, несовместимых с ценностями рассматриваемой группы. Само существование Коллективного Антагониста угрожа-

ет группе, и его ликвидация рассматривается как одна из целей, которые должны быть достигнуты в ходе деятельности группы.

Реализация такой цели возможна при дегуманизации коллективного антагониста. Наиболее распространённые методы дегуманизации оппонента в интернет-сообществах — использование устойчивых «штампов-обзывалок» для носителей противоположной идентичности, постулирование принципиальной невозможности диалога с ними. Штампы-названия, унижающие оппонента по признаку его принадлежности к группе, всегда являются составной частью единой идеологии, переосмысливающей как видение истории, так и религиозные, идеологические, культурные принципы, бытующие в обществе. Самыми важными из таких штампов являются этнофолизмы и религиофолизмы, то есть экспрессивные пейоративные термины с резко негативной коннотацией, унижающие антагониста по этническому или религиозному принципу (подробнее см. [6]).

4.2. Реальные экстремистские группы, формирующие коллективную элитарность с использованием терминов-антагонистов

На данный момент в русскоязычном Интернете распространяются две экстремистские идеологии, носители которых проявили себя в действии, проводя террористические акты, участвуя в военных действиях, совершая военные преступления. Это радикальная мусульманская идеология ИГИЛ и украинская националистическая идеология. Термины-антагонисты, используемые в этих идеологических системах: религиофолизм «кафиры» (ИГИЛ) и этнофолизмы «москали/ватники/вата» (украинские националисты).

Итак, в двух реальных экстремистских группах для называния Коллективного Антагониста используются религиофолизм «кафиры» и этнофолизм «ватники» соответственно.

В работе с помощью системы IQBuzz исследовались две коллекции документов из различных социальных сетей, собранные по запросам «кафиры» и «ватники».

При исследовании коллекции документов было выявлено несколько тем из интересующего нас предметного поля, а именно: описание собственной группы; описание Коллективного Антагониста; война или борьба между ними.

Для описания собственной группы используются слова, представляющие своих сторонников как борцов за правду и добро, а именно: военные, бойцы, солдаты, джихад, включая слова, описывающие действия собственной группы: убиты, ранены, погребены, восстание, уничтожили, сражаются.

В описании Коллективного Антагониста используется экспрессивная лексика, негативно окрашенная. Действия антагониста вписываются в схему военного противостояния. К этой теме относятся термины террористы, националисты, ватники, банды, вооруженных, тупые, оккупанты, москаль, предатели, каратели, кафиры, свиньи, рабы, фанатики, лицемеры, наемники, а также слова, описывающие действия противника: резать, геноцид, напали, трупы, были убиты, бомбят, сожгли, невинных, нагло, незаконно, убивать, продали, кровавый, зверства.

Таким образом, наиболее перспективными для исследования представляются термины, которые используются для названия Коллективного Антагониста (термины-антагонисты): они эмоционально нагружены, очерчивают границы идентичности, проводя линию свой-чужой, и встречаются в текстах экстремистской направленности.

4.3. Анализ терминов в двух коллекциях документов, отобранных по критерию использования терминов-антагонистов

В ходе работы были проанализированы с помощью программы IQbuzz две коллекции документов. Принцип формирова-

ния коллекций: употребление терминов-антагонистов, сформированных в двух экстремистских средах, а именно в радикальной мусульманской и украинской националистической. Одна коллекция составлена из всех документов, где употребляется слово кафир и его падежные формы, во второй коллекции используются все формы слов ватник и москаль.

Для каждой коллекции был составлен частотный словарь и выделена общая лексика, то есть общие термины, которые есть в обеих коллекциях. Также было зафиксировано количество вхождений каждого общего термина в третью большую коллекцию по посторонней теме.

Для определения источников экстремистской идеологии использовался разработанный авторами ИИВ, который зависит от количества неявных связей между сайтами и строится на основе лингвистического процессора (ЛПНС), выделяющего неявные связи между текстами с помощью меры семантического подобия. Апробация методики на конкретных примерах — анализ коллекций на тему «ватники-москаль», «кафиры» — подробно изложены в статьях [6, 16].

Так, термин кафиры входит в первую коллекцию 2951 раз, во вторую 5 раз, в коллекции из других предметных областей не встречается ни разу. Термин москаль тоже не встречается в коллекции документов на посторонние темы, в первой коллекции встречается 3 раза, во второй — 732 раза.

С помощью ИИВ были найдены наиболее значимые статьи (например, в Вики), содержащие много ключевых фраз («кто не скачет — тот москаль» и др), которые позволяют проследить источники и наиболее влиятельные документы, содержащие экстремистскую идеологию.

Таким образом, использование ИИВ позволяет выявлять тексты-первоисточники различных идей в социальных сетях и в Интернете, позволяет прослеживать потоки распространения идей, а также быстрее

находить достоверную и качественную информацию из первоисточников. Анализ полученных таким образом коллекций позволяет выделить разные группы характерных терминов.

4.4. Выделение разных групп характерных терминов

Работа с выборкой терминов в двух коллекциях, сформированных с помощью ИИВ, позволяет разделить их на группы.

К первой группе относятся термины из трех указанных выше тем (война; описание своей группы; описание Коллективного Антагониста), которые входят в общие отобранные коллекции примерно одинаковое число раз. Это общие характерные термины (ОХТ). Анализируя эти термины, можно выделить следующие темы, общие для обеих коллекций:

1. тема война представлена терминами
 - 913793710779 9989 ##### k=убиты
v=369 V=326 Vm=0
 - 970903388739 9981 ##### k=сражались
v=202 V=201 Vm=0
 - 984095113779 9983 ##### k=невинных
v=137 V=162 Vm=0
2. тема терроризм представлена терминами
 - 981919425099 9985 ##### k=теракты
v=145 V=174 Vm=0
 - 983345518799 9981 ##### k=терроризм
v=166 V=140 Vm=0
 - 996376718719 9977 ##### k=экстремисты
v=64 V=79 Vm=0
 - 996881226239 9985 ##### k=казнили
v=64 V=68 Vm=0
3. призывы к насилию
 - 997678118799 9979 ##### k=устрашения
v=36 V=90 Vm=0
 - 997903857249 9981 ##### k=вероломно
v=45 V=65 Vm=0
 - 997838643919 9985 ##### k=святыни
v=52 V=58 Vm=0
 - 997547692139 9983 ##### k=взрывают
v=59 V=58 Vm=0

998674951129 9973 ##### k=возненавидели
v=43 V=43 Vm=0

Следующая группа, выделенная при работе с двумя коллекциями — специфические характерные термины (СХТ), которые чаще встречаются в документах одной из групп.

В документах с мусульманской радикальной идеологией специфические характерные термины формируют следующие темы:

1. священная война — представлена терминами
 - 950932343899 9987 ##### k=джихад
v=1670 V=41 Vm=0
 - 962928730099 9975 ##### k=мусульманами
v=1478 V=35 Vm=0
2. оскорбления антагонистов с использованием религиофилемов:
 - 965333023749 9983 ##### k=неверных
v=1075 V=45 Vm=0
 - 996134497779 9983 ##### k=кафирами
v=1798 V=3 Vm=0
 - 996481346699 9987 ##### k=кяфиры
v=982 V=5 Vm=0

В коллекции, выражающей украинскую националистическую идеологию, специфические характерные термины формируют следующие темы:

1. национализм — представлена терминами
 - 915975132499 9973 ##### k=националистов
v=70 V=1675 Vm=0
 - 943793275839 9975 ##### k=националисты
v=57 V=1376 Vm=0
 - 949110670439 9975 ##### k=национализма
v=82 V=866 Vm=0
 - 935296910559 9975 ##### k=политический
v=99 V=912 Vm=0
2. Великая Отечественная война
 - 974631297999 9983 ##### k=фашистов
v=30 V=1180 Vm=0
 - 998007051969 9981 ##### k=Оккупанты
v=9 V=309 Vm=0
 - 997989852849 9985 ##### k=палачей

v=17 V=165 Vm=0
 955174793499 9985 ##### k=геноцид
 v=90 V=695 Vm=0
 993624859519 9985 ##### k=Бандера
 v=8 V=1112 Vm=0

3. оскорбления антагонистов по политическому признаку:
 975600898389 9975 ##### k=сепаратистов v=39 V=873 Vm=0
 935296910559 9975 ##### k=политический v=99 V=912 Vm=0
 998084448009 9977 ##### k=референдуме v=11 V=243 Vm=0
 998680325854 9979 ##### k=пропаганда v=127 V=1450 Vm=10

Общие характерные термины относятся к темам войны и терроризма. Можно предположить, что одни и те же события описываются одной стороной как война, которую ведут бойцы, а другой — как серия терактов, совершаемых террористами. Такая двойная структура выражает идейные основы существования экстремистских групп с коллективной элитарностью, а частое употребление экспрессивной терминологии в этой теме указывает на основные направления деятельности группы, которые требуют эмоциональной вовлеченности носителей групповой идентичности.

Специфические термины, которые выходят на первые места в своих группах, формируют особые темы, характерные только для данной группы. Для первой группы это джихад, мусульмане, муслимы, неверующий. Для второй — националисты, сепаратисты, фашисты, пропаганда, геноцид. В первой группе идейно нагруженными оказываются термины религиозного значения, во второй — термины националистической семантики и термины, связанные с Великой Отечественной войной 1941 - 1945 гг.

Таким образом, проведенный методами тематического моделирования анализ подтверждает распространение в сети экстремистских идеологий, основным признаком ко-

торых - формирование коллективной элитарности, причем критерий такого формирования является специфическим для каждой идеологии. Идеология мусульманских экстремистов имеет религиозную основу, идеология украинских националистов основана на политической националистической концепции и включает в себя переосмысление событий Великой Отечественной войны в нацистском ключе.

4.5. Динамика терминологии и её роль в прогнозировании тревожных событий.

Работа с терминологией корпусов проводилась в течение трех лет. Каждый документ в IQbuzz снабжается датой появления в сети, что позволяет проследить динамику терминологии и выявить некоторые закономерности.

1. Наблюдается рост корпусов текстов, содержащих экстремистскую идеологию, как и рост количества сайтов, такую идеологию распространяющих.
2. С ростом числа текстов растёт количество общих характерных терминов в обоих корпусах.
3. Для каждого из терминов технически возможно проследить этапы количественного роста или спада.
4. Эмоциональный накал в теме меняется в зависимости от этапа роста или спада в употреблении термина.
5. Эмоциональная вовлечённость читателя максимальна в период роста, снижается в период спада.
6. Вероятность тревожных событий в период роста повышается, в период спада понижается.

Проиллюстрируем сказанное данными для термина «кафир» [17]. Это термин-антагонист, по которому выделялась группа, распространяющая радикальную мусульманскую идеологию в сети (рис. 2).

Виден рост в 2015 и 2016 годах. Спад в 2017 году коррелирует с разгромом ИГИЛ в Си-



Рис. 2: Количество записей в социальных сетях с термином-антагонистом «Кафир» в день (с 2014-01-01 по 2017-03-01)

рии в результате действий правительственных войск, поддержанных РФ.

На основании проведённых исследований можно высказать следующие гипотезы:

1. Рост общей характерной терминологии двух коллекций происходит за счёт усиления жесточённости обеих групп терминов, то есть использования самых эмоционально нагруженных слов с негативной семантикой, выражающих максимальную жесточённость.
2. Если разбить ОХТ на группы по степени жесточённости, можно выявить корреляции с этапами количественного роста/спада терминов разных тем.
3. Выявленные при исследовании закономерности позволяют измерять степень жесточённости и как следствие степень экстремизма терминологии, определять этап количественного роста/спада и прогнозировать будущие этапы.
4. Прогноз этапов количественного роста/спада экстремистской терминологии в текстах, продуцируемых найденными сайтами, позволяет определить степень эмоциональной вовлечённости в какую-либо тему и спрогнозировать рост или спад интереса к этой теме, который будет реализован в росте или спаде количества тревожных событий в реальности.

5. Заключение и план дальнейших исследований

В ходе работы авторским коллективом опробована методика тематического моде-

лирования с использованием ИИВ для нахождения сайтов распространяющих экстремистскую идеологию. Найдены динамические закономерности употребления терминов разных тем и обнаружена корреляция этой динамики с реальными тревожными событиями

Для получения результата сформированы две коллекции документов реально существующих групп с экстремистской идеологией, а именно радикальных мусульман и украинских националистов. Экстремистская идеология обеих групп основана на формировании коллективной элитарности. Семантически это выражается в использовании терминов-антагонистов — эмоционально негативно окрашенных религио- и этнофолизмов. Использование терминов-антагонистов стало основанием, по которому были сформированы две коллекции документов с помощью программы IQbuzz. Далее были выделены общие характерные термины для обеих коллекций и специфические характерные термины, преобладающие в каждой из коллекций. При анализе терминов выделены темы, которые указывают на деятельность группы по следующим направлениям: формирование коллективной элитарности, оскорбление антагонистов, призывы к насилию по отношению к ним, представление своей деятельности как справедливой войны за высшие идеалы - религиозные и политические соответственно.

Дальнейший план работ включает анализ многоязычных коллекций; поиск новых экстремистских групп по указанным темам; поиск наиболее значимых текстов в выявленных группах с использованием ИИВ; анализ динамики роста тем в коллекциях по месяцам. Выполнение этих планов позволит находить экстремистские группы в сети; выделять самые влиятельные тексты, которые наиболее полно выражают основные темы новой экстремистской идеологии; рассчитывать динамику распространения основных тем и на этой основе прогнозировать тревожные события.

Литература

- [1] Nockleby John T. (2000), "Hate Speech" in Encyclopedia of the American Constitution, ed. Leonard W. Levy and Kenneth L. Karst, vol. 3. (2nd ed.), Detroit: Macmillan Reference US, pp. 1277–79. Cited in "Library 2.0 and the Problem of Hate Speech," by Margaret Brown-Sica and Jeffrey Beall, Electronic Journal of Academic and Special Librarianship, vol. 9 no. 2 (Summer 2008).
- [2] Bettencourt L. et al. The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models // Physica A. — 2006. — May. — Pp. 513–536.
- [3] Jacobs J.A. Receptivity Curves: Educational Research and the Flow of Ideas: Expanded Version, Population Studies Center, University of Pennsylvania, PSC Working Paper Series, PSC 13–10. — 2013. — http://repository.upenn.edu/psc_working_papers/50
- [4] Jee J., Klippel, L., Hossain, M., Ramakrishnan Ramakrishnan, N., Mishra, B. Discovering the Ebb and Flow of Ideas from Text Corpora // IEEE Computer 45(2). — 2012. — Pp. 73–77.
- [5] Золотарев О.В., Шарнин М.М., Клименко С.В. Семантический подход к анализу террористической активности в сети Интернет на основе методов тематического моделирования. // Вестник Российского нового университета. Серия: Сложные системы: модели, анализ и управление. М.: РосНОУ, 2016. № 3. С. 64-71.
- [6] Клименко С.В., Шарнин М.М., Ищенко Н.С., Хакимова А.Х. Использование метода неявных ссылок для противодействия распространению деструктивных антигуманистических идеологий в сети Интернет // Международная Школа-семинар «Ситуационные центры и информационно-аналитические системы класса 4i для задач мониторинга и безопасности» SCVRT2015-16, Пущино, Царьград, 21-24 ноября 2015-2016 гг., С. 129-133.
- [7] Меркурьева Н.А., Шарнин М.М., Слободюк Е.А., Цыганов В.В. Построение визуальных карт предметной области для разведочного поиска (exploratory search) // Сборник трудов Международной конференции СРТ2017, 07-14 мая 2017, Ларнака, Кипр. Изд-во Института физико-технической информатики (ИФТИ). 2017. С.271-274.
- [8] Цыганов В.В., Шарнин М.М., Родина И.В. Построение визуальных карт предметной области с помощью методов тематического моделирования по данным социальных сетей. // Сборник трудов Международной конференции СРТ2017, 07-14 мая 2017, Ларнака, Кипр, Изд-во Института физико-технической информатики (ИФТИ). 2017. С.355-359.
- [9] Michael Charnine, Stanislav Klimenko. Semantic cyberspace of scientific papers. // 2017 International Conference on CYBERWORLDS. University of Chester, Chester, United Kingdom. 20-22nd September, 2017. In cooperation with the Eurographics Association, ACM SIGGRAPH, and the International Federation for Information Processing (IFIP) — <http://www.cyberworlds-conference.org/> (в печати).
- [10] Steyvers, M., Griffiths, T. Probabilistic Topic Models. — Hillsdale, NJ, USA : Erlbaum, 2007.
- [11] IQBuzz.pro — Сервис мониторинга социальных медиа и онлайн СМИ.// <http://iqbuzz.pro/about.php>
- [12] Galina I., Charnine M., Somin N., Nikolaev V., Morozova Yu., Zolotarev O.. "Metod for generating subject area associative portraits: different examples". // Proceedings of the 2015 International Conference on Artificial Intelligence (ICAI 2015), WORLDCOMP'15, July 27-30, 2015. Las Vegas Nevada, USA, v.I, pp.288-294.
- [13] Галина И.В., Козеренко Е.Б., Морозова Ю.И., Сомин Н.В., Шарнин М.М. Ассоциативные портреты предметной области (аппо) — инструмент автоматизированного построения систем big data для извлечения знаний: теория, методика, визуализация аппо, возможное применение. // Информатика и её применения. 2015. Т.9. Вып.2. С.94-111.
- [14] Charnine Michael, "Keywen: Automated Writing Tools", Booktango, USA, 2013, ISBN 978-1-46892-205-9.
- [15] Шарнин М. М., Сомин Н. В., Кузнецов И. П., Морозова Ю. И., Галина И. В., Козеренко Е. Б.. Статистические механизмы формирования ассоциативных портретов предметных областей на основе естественно-языковых текстов больших объемов для систем извлечения знаний. // Информатика и её применения. 2013. Т.7. №2. Стр.92–99.
- [16] Хакимова А.Х., Шарнин М.М., Клименко С.В., Золотарев О.В., Родина И.В. Мера подобия текстов как инструмент оценки интертекстуальности при анализе больших коллекций документов. // Вестник Российского нового университета, М.: РосНОУ, 2016. — Выпуск 4, С.62-71.
- [17] Клименко С.В., Шарнин М.М., Галина И.В., Мацкевич А.Г., Хакимова А.Х., Соколов Е.Г. Индекс идеологического влияния: базовые модули комплексной методики.

Современные подходы к обработке многоязычных текстов, основанные на методах дистрибутивной семантики*

О. В. Золотарев¹, М. М. Шарнин², Алена Еромасова¹, Ф. М. Тезадова¹
ol-zolot@yandex.ru

Москва, ¹АНО ВО «Российский новый университет»

Москва, ²Федеральный исследовательский центр «Информатика и управление» РАН

Рассматриваются вопросы обработки текстов естественного языка с использованием методов дистрибутивной семантики. В работе описываются методы статистической обработки естественно-языковых текстов, основанных на векторном представлении информации, приведены примеры реализации данных методов при обработке больших корпусов документов (Big Data).

Ключевые слова: word2vec, векторное представление слов, дистрибутивная семантика, многоязычная семантическая сеть, BabelNet, WMD.

Geotectonics, as a reason for the difference in the results of a comparative assessment of internal and external exposure and damage to the population from nuclear power plants*

О. В. Золотарев¹, М. М. Шарнин², Алена Еромасова¹, Ф. М. Тезадова¹

Moscow, ¹Russian New University

²FRC CSC RAS

The problems of processing natural language texts using the methods of distributive semantics are considered. The paper describes methods of statistical processing of natural language texts based on the vector representation of information, gives examples of implementing these methods for processing large document bodies (Big Data).

Keywords: word2vec, vector representation of words, distributive semantics, multilingual lexicalized semantic network BabelNet, WMD.

Введение

В последнее время все большее значение приобретают вопросы анализа информации, как с точки зрения обеспечения безопасности, так и с точки зрения исследования все возрастающих объемов текстов в сети Интернет для извлечения ценной информации и знаний [1]. В данной статье рассматриваются вопросы автоматической обработки больших объемов информации, представленной в электронном виде.

В данной статье рассматриваются подходы к обработке естественно-языковых массивов информации: word2vec, WMD, а также многоязычная энциклопедия, онтология и одновременно семантическая сеть BabelNet по представлению знаний в сети Интернет. Word2vec представляет собой технологию для анализа семантики естественных языков [2]. В результате обработки кор-

пусов текстов строится матрица векторов слов. WMD — методика сравнения текстов и одновременно технология, основанная на word2vec.

Все упомянутые выше методы предполагают проведение предварительной обработки текстов [3, 4, 5]. Сначала выполняется удаление так называемых стоп слов — предлогов, союзов, артиклей и т. д. Затем происходит нормализация слов (единственное число, именительный падеж — для существительных, неопределенная форма — для глаголов и т. д.), удаляются слишком часто и очень редко встречающиеся слова. Далее выделяются концепты, формируются синсеты, определяются синонимические и семантические связи [6, 7].

Word2vec

Данная технология применяется для выявления закономерностей в рамках контекстного окружения слов путем построения матрицы их совместной встречаемости.

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 16-07-00756, 16-29-09527, 18-07-00909, 18-07-01111, 18-07-20030.

В такой матрице значения слов представляются в виде семантических векторов [8].

Алгоритм работы Word2Vec [2]:

1. Составление коллекции текстов для обучения нейронной сети.
2. Составление словаря слов коллекции с подсчетом частоты встречаемости каждого значащего слова. В словарь не попадают стоп-слова (местоимения, частицы и т. д., а также часто встречающиеся слова).
3. Установка размерности векторов (300 - 500). Элементы вектора — это пары: признаки и веса (признак — значимое слово коллекции; вес — количество совместной встречаемости главного слова со словом окрестности — см. далее). Иногда назначают случайные значения элементам вектора в качестве начальных условий. Но лучше устанавливать значения элементов в ноль.
4. Выбор размера скользящего окна (3 - 10).
5. Первый проход по коллекции текстов со скользящим окном, определение в рамках окна главного слова — для которого будет строиться вектор. Если размер скользящего окна равен 3, то главное слово — второе. Анализируется его окружение — одно слово вправо, одно слово влево.
6. Вычисление частоты попарной встречаемости с главным словом окружающих слов в окрестности скользящего окна. Если в рамках текущего положения окна для главного слова встречается любое другое слово из окрестности более одного раза, это слово в векторе учитывается только один раз для данного положения скользящего окна. Каждый раз при встрече слова окрестности с главным словом значение веса слова окрестности увеличивается на 1.
7. Второй проход по коллекции текстов. Вычисление векторов слов. Для каждого слова в результате первого прохода мы имеем:

- (а) Количество раз, которое данное слово встретилось в коллекции текстов.
- (б) Для каждого слова окрестности главного слова подсчитано количество раз, которое слово окрестности находилось в окрестности главного слова.
Далее вычисляются веса элементов (слов окрестности) вектора главного слова. Может быть предложено большое количество методов для расчета весов, например, значение веса каждого слова делится на количество частоту встречаемости главного слова в коллекции. Получаем значения весов (элементов вектора) от 0 до 1 для каждого слова в коллекции.

Построенная матрица векторов представляет собой нейронную сеть с описание выявленных закономерностей предметной области в рамках корпуса текстов на основе совместной встречаемости слов. При поступлении нового текста на вход нейронной сети она может либо достраиваться за счет пересчета частот совместной встречаемости слов, либо на основе сравнения матриц векторов может делаться заключение о степени близости данного текста предметной области.

Над векторами возможно выполнение математических операций сложения и вычитания для определения близости понятий и предсказания результата.

Определение степени семантической близости слов в векторном пространстве может выполняться посредством расчета косинусного расстояния между векторами. С помощью этого метода измеряется косинус угла между векторами слов, который является показателем их смысловой схожести

$$\cos(w_1, w_2) = \frac{\vec{V}(w_1) \times \vec{V}(w_2)}{|\vec{V}(w_1)| \times |\vec{V}(w_2)|}. \quad (1)$$

В числителе содержится скалярное произведение векторов слов w_1 и w_2 :

$$\vec{V}(w_1) \times \vec{V}(w_2) = \sum_{i=1}^n w_{1i}w_{2i}. \quad (2)$$

В знаменателе содержится произведение евклидовых норм этих векторов:

$$|\vec{V}(w)|_B = \sqrt{\sum_{i=1}^n |w_i|^2}. \quad (3)$$

При поступлении неизвестного текста на вход нейронной сети для анализа могут использоваться 2 алгоритма для предсказания либо слова по его окружению (метод CBOW), либо окружения слова по его контексту (метод Skip-gram) [2].

Сферы применения Word2vec:

- выявление семантической близости слов;
- кластеризация слов по семантической близости;
- оптимизация поисковых запросов: чтобы находились документы/сайты не только с конкретными словом и с четкими синонимами, но и с нечеткими синонимами и тематически ассоциированными словами;
- машинный перевод;
- выявление неявных ссылок между документами;
- поиск схожих по смыслу документов;
- антиплагиат и т. д.

WMD — Word Mover's Distance

WMD — один из новых методов дистрибутивной семантики, который позволяет определять различие между двумя текстовыми документами путем вычисления минимального расстояния между векторами слов каждого документа. Основой для работы данного метода служит метод word2vec. Вычисление минимального расстояния между векторами слов может осуществляться, как уже было сказано выше, за счет операций над векторами (вычитание, скалярное произведение) [9].

Сферы применения WMD:

- выявление именованных сущностей;
- тэгирование частей речи;
- машинный перевод;
- кластеризация документов;
- ранжирование документов;
- анализ тональности текста.

BabelNet

Основными элементами BabelNet являются синонимические единицы — синсеты, которые представляют собой некоторые понятия предметной области, связанные разного рода отношениями [10]. Онтология может формироваться либо вручную, либо автоматически с использованием методов дистрибутивной семантики [11].

Для построения онтологии используются корпуса текстов предметной области. Процесс построения онтологии (семантической сети) представлен на рисунке 1.

На сегодняшний день BabelNet содержит знания, представленные более чем на 280 языках. Это говорить о том, что данная сеть является общепризнанным информационным ресурсом. Данная сеть объединила в себе более 40 разнородных ресурсов, включая Википедию, WordNet, Wikidata, Викисловарь, OmegaWiki, VerbNet, Microsoft Terminology, Wikiquote, GeoNames, WoNeF, ItalWordNet, FrameNet, Open Dutch WordNet и ImageNet.

Интеграция данных ресурсов выполняется с помощью автоматического алгоритма связывания и путем заполнения лексических зазоров с помощью машинного перевода. Результатом интеграции является энциклопедический словарь, который содержит Babel-синсеты, представленные на множестве языков и связанные огромным количеством семантических отношений: в BabelNet 4.0 из 832 млн значений выделено более 6 млн понятий, более 9,5 млн именованных сущностей, связанных между собой более чем 1 млрд семантических связей.

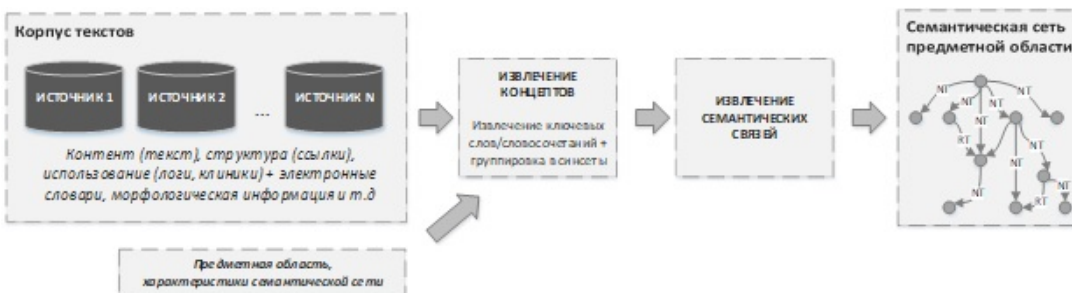


Рис. 1: Процесс построения онтологии (семантической сети)

Структура BabelNet:

1. Отдельные понятия или именованные сущности (персоны, продукты, организации и т. д.).
2. Синсеты, описывающие в том числе понятия, могут подразделяться на части речи (существительное, прилагательное, глагол, наречие).
3. Семантические отношения: мероним/холоним, антоним и т. д.
4. Ситуации, в которых встречаются синсеты (в зависимости от категории синсетов).
5. Ключевые понятия, которые используются для связи понятий из разных источников.

Сферы применения BabelNet:

- обработка естественного языка;
- перевод;
- текстовый анализ;
- классификация документов;
- антиплагиат;
- кластеризация документов;
- поиск документов по схожей тематике.

Благодарности

Работа выполнена при поддержке Российского фонда фундаментальных исследований, гранты 16-07-00756, 16-29-09527, 18-07-00909 и 18-07-01111. Мы благодарны Российскому фонду фундаментальных исследований за финансовую поддержку наших проектов.

Заключение

Как следует из выше описанного сфера использования интеллектуальных инструмен-

тов анализа текста достаточно обширна и постоянно расширяется. В статье рассмотрены технологии автоматической обработки текста, основанные на статистике.

В связи с быстрорастущим объемом информации в сети Интернет использование единой интегрированной среды, объединяющей разнородные ресурсы для обмена информацией, позволит в существенной степени упростить межязыковые, межкультурные коммуникации. BabelNet может использоваться и в исследовательских целях. Важно, что ресурс остается открытым для доступа, постоянно развивается, совершенствуется. Говоря о семантическом анализе, нельзя сказать, что описанные технологии выявляют смысл слов, фраз, текста, но они могут помочь выделить характерные концепты, их связи, определить их сходство с достаточно высокой степенью вероятности.

Описанные подходы позволят не только проводить анализ документов на антиплагиат посредством сравнения текстов документов, но и с учетом более глубоких синонимических связей [12].

Литература

- [1] Шарнин М.М., Золотарев О.В., Сомин Н.В. Извлечение и обработка знаний из неструктурированных текстов деловой сферы и социальных сетей. В сборнике: Социальный компьютер: основы, технологии развития, социально-гуманитарные эффекты. Материалы Четвертой Международной научно-практической конференции. 2015. С. 364-371.
- [2] Mikolov T., Corrado G., Кю Chen, Лю Dean, Efficient Estimation of Word Representations in Vector Space, Proceedings of the International Conference on Learning Representations (ICLR 2013), 1, 2013, 1-12.

- [3] Zolotarev O.V. , Charnine M.M., Matskevich A.G., Kuznetsov K.I. Business Intelligence Processing on the Base of Unstructured Information Analysis from Different Sources Including Mass Media and Internet. Proceedings of the 2015 International Conference on Artificial Intelligence (ICAI 2015), vol.I, WORLDCOMP'15, July, 27-30, 2015, Las Vegas Nevada, USA, I, 295-299.
- [4] Galina I.V., Charnine M.M., Somin N.V., Nikolaev V.G., Morozova Y.I., Zolotarev O.V. Method for Generating Subject Area Associative Portraits: different Examples. Proceedings of the 2015 International Conference on Artificial Intelligence (ICAI 2015), I, WORLDCOMP'15, July 27-30, 2015. Las Vegas Nevada, USA, I, 288-294.
- [5] Zolotarev O.V., Charnine M.M., Matskevich A.A. A Conceptual Business Process Structuring by Extracting Knowledge from Natural Language Texts. Proceedings of the 2014 International Conference on Artificial Intelligence (ICAI 2014), I, WORLDCOMP'14, July 21-24, 2014, Las Vegas Nevada, USA. CSREA Press, 82-87.
- [6] Abdelwahab O., Elmaghraby A. UoFL at SemEval-2016 Task 4: Multi Domain word2vec for Twitter Sentiment Classification, SemEval, 2016, 169-175.
- [7] О.В. Золотарев Методы выделения процессов, объектов, отношений из текстов естественного языка. Проблемы безопасности российского общества. 2014. № 3-4. С. 276-283.
- [8] Tian J., Lan M. ECNU at SemEval-2016 Task 1: Leveraging Word Embedding from Macro and Micro Views to Boost Performance for Semantic Textual Similarity, SemEval-2016, 618-624.
- [9] Kusner M. J., Sun Y., Kolkin N. I., and Weinberger, K. Q. From word embeddings to document distances. In ICML, 2015.
- [10] Roberto Navigli and Simone Paolo Ponzetto. 2012. BabelNetXplorer: a platform for multilingual lexical knowledge base access and exploration. In Comp. Vol. to Proc. of WWW-12, pages 393–396.
- [11] Клименко С.В., Золотарев О.В., Шарнин М.М. Использование онтологического подхода для анализа текстов естественного языка. Вестник Российского нового университета. Серия: Сложные системы: модели, анализ и управление. 2017.№ 1. С. 67-71.
- [12] Михеев М.Ю., Сомин Н.В., Галина И.В., Золотарев О.В., Козеренко Е.Б., Морозова Ю.И., Шарнин М.М. Фальштексты: классификация и методы опознания текстовых имитаций и документов с подменой авторства // Информатика и ее применения. Том 8, выпуск 4, РАН, — М.: 2014, с. 70-77.

Методы оценки семантического сходства как средства обнаружения межъязыкового плагиата*

А. Х. Хакимова¹, М. М. Шарнин², К. И. Кузнецов², Ю. В. Сюракшина³
aida_khatif@mail.ru, mc@keywen.com, k.smith@mail.ru,
yuliya.surakshina@phystech.edu

Протвино, ¹АНО «Институт физико-технической информатики»

Москва, ²Институт Проблем Информатики ФИЦ ИУ РАН

³Московский физико-технический институт (государственный университет)

Приведен обзор методов оценки семантического сходства разноязычных текстов, применяемых для обнаружения межъязыкового плагиата. Выделено пять семейств поисковых моделей, основанных на: синтаксисе, тезаурусе, сопоставимых корпусах, параллельных корпусах и машинном переводе. Дано формальное описание задачи автоматического обнаружения плагиата. Приведен пример авторского алгоритма и программы оценки межъязыкового семантического сходства для текстов на русском, украинском и английском языках.

Ключевые слова: межъязыковой плагиат, обнаружение плагиата, семантическое сходство, семантический вектор, мегалеммы, автоматическое обнаружение плагиата, языковые пары

The methods for estimation of semantic similarity as a means of cross-language plagiarism detection*

A. K. Khakimova¹, M. M. Charnine², K. I. Kuznetsov², Y. V. Surakshina³

Protvino, ¹ANO "Institute of computing for physics and technology"

Moscow, ²FRC CSC RAS

³Moscow Institute of Physics and Technology (State University)

A review of methods for assessing the semantic similarity of multilingual texts used to detect cross-language plagiarism is presented. Five families of search models based on syntax, dictionary, comparable corpora, parallel corpora and machine translation are identified. A formal description of the automatic detection of plagiarism is given. An example of an author's algorithm and a program for evaluating cross-language semantic similarity for texts in Russian, Ukrainian and English is presented.

Keywords: cross-language plagiarism, detection of plagiarism, semantic similarity, semantic vector, megalmmas, automatic detection of plagiarism, language pairs

Введение

К настоящему времени в мире идет усиление борьбы с неправомерным использованием чужой интеллектуальной собственности. В случае письменного языка плагиат означает включение в документ фрагментов текста, написанного кем-то другим, без предоставления соответствующей ссылки.

В наши дни проблемы, связанные с плагиатом и, в частности, плагиатом текстов, возросли благодаря простому доступу к крупным источникам информации с помощью электронных средств, таких, как электронные библиотеки и, конечно же, Интернет. Количество случаев плагиата документов

(либо полного, либо фрагментарного) за последние несколько лет значительно возросло. Одной из основных причин развития этого явления, как отмечалось выше, является легкость, с которой можно получить доступ к документам в электронной поддержке.

К сожалению, обнаружение плагиата вручную практически невозможно. Следовательно, важно разработать автоматизированные механизмы, позволяющие выполнять задачу обнаружения плагиата в разумные сроки.

Плагиат разделяют на различные типы, такие, как буквальный, интегральный, внутренний, внешний, точная копия, манипуляции с текстом и т. д. [1]. Подобным же об-

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 16-07-00756, 16-29-09527, 18-07-00909, 18-07-01111, 18-07-20030, 18-37-10008.

разом существуют различные методы для обнаружения плагиата.

Согласно [2], **межъязыковое обнаружение плагиата (CLPD)** заключается в распознавании семантически подобных текстов независимо от языков, на которых они написаны, когда не приводится ссылка на исходный источник.

В одноязычном контексте проблема обнаружения плагиата хорошо разработана недавними исследованиями [3]. Тем не менее, расширение Интернета облегчает доступ к документам во всем мире, а свободно доступные инструменты машинного перевода способствуют распространению межъязыкового плагиата.

Межъязыковой плагиат означает плагиат путем перевода вручную или автоматически. Проблема в обнаружении такого рода плагиата заключается в том, что подозрительный документ больше не находится в языковом поле его источника.

В этой относительно новой области исследований было проведено определение семантического текстового подобия на языковых парах. Авторы исследовали различные существующие подходы обнаружения плагиата на разных языковых парах и обнаружили, что если метод эффективен для конкретной языковой пары, он будет таким же эффективным для другой языковой пары при достаточном количестве доступных лексических ресурсов, т.е. метод может быть оптимизирован для конкретного корпуса и эффективно применен на другом корпусе [5].

Современные методы обнаружения межъязыкового плагиата

Определение межъязыкового семантического текстового подобия является основной частью процесса обнаружения межъязыкового плагиата.

На данный момент выделено пять типов подходов для обнаружения межъязыкового плагиата. Цель каждого метода — оце-

нить, выражают ли две текстовые единицы на разных языках одно и то же сообщение, или нет. На рис. 1 представлена таксономия Potthast et al. (2011) [6], обогащенная Danilova (2013) [7], включающая методы обнаружения межъязыкового плагиата, сгруппированные по классам подходов.

Однако стоит отметить, что термин «сходство» здесь означает, что объекты (тексты) разделяют только некоторые характеристики и сопоставимы, тогда как плагиат имеет отношение к случаям, когда авторские слова и идеи копируются (с формальными изменениями или без них). Как следует из обновленной версии определения плагиата Barrón-Cedeño (2012) [8], — пример межъязыкового плагиата происходит, когда мы имеем дело с неподтвержденным повторным использованием текста, переведенного с одного языка на другой.

Вопросы обнаружения межъязыкового плагиата стали привлекать исследователей с момента объявления **Международного Конкурса на обнаружение плагиата PAN (Uncovering Plagiarism, Authorship and Social Software Misuse)** в 2009 г., вопросы межъязыкового семантического сходства начали привлекать внимание исследователей.

В 2010 году были попытки использования *машинного перевода* (МТ) в качестве шага предварительной обработки документа, чтобы иметь дело с неанглийскими документами как возможными источниками плагиата. Основными проблемами, которые немедленно проявились, были вычислительные затраты и низкое качество МТ, которое не способно сравнивать подозрительные тексты и источники. Более того, авторы склонны модифицировать переведенные тексты, используя парафраз, что делает процесс распознавания еще сложнее. Кроме того, одной из основных проблем является наличие значительных различий в синтаксических структурах языков, принадлежащих к разным семьям. Несмотря на недостатки подхода, основанного на МТ, исследователи от него не отказались. Как

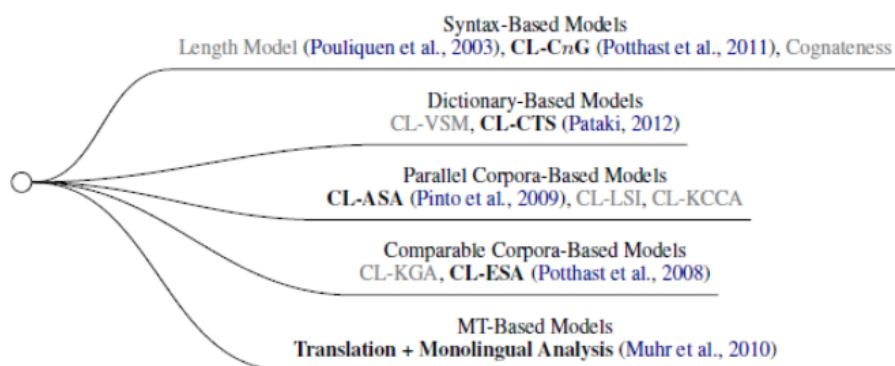


Рис. 1: Таксономия Potthast et al. (2011) [6], уточненная Danilova (2013) [7], различных подходов к обнаружению межъязыкового семантического сходства.

отмечают Meuschke и Gipp (2013) [9], он подходит для небольших коллекций документов.

Исследования Potthast et al. (2011) [6] и Barrón-Cedeño et al. (2013) [10] были посвящены исключительно классификации и оценке методов *CLPD*.

Barrón-Cedeño (2012) [8] делит методы *CLPD* на внутренние и внешние, поскольку, как показано в литературе, методы обнаружения встроенного плагиата позволяют различать так называемые эффекты процесса перевода внутри текста. Некоторые из соответствующих индикаторов, найденных исследователями, следующие: функциональные слова, морфосинтетические категории, личные местоимения, наречия, анимированные местоимения (такие, как я, мы, он), связные маркеры. Некоторые исследователи, процитированные Pataki (2012) [11], утверждают, что в результате серии экспериментов с немецко-английским переводом никаких закономерностей, указывающих на МТ, в текстах не было выявлено.

Большинство авторов связывают *CLPD* с внешним обнаружением плагиата, следовательно, процесс включает такие же шаги, как и при обычном обнаружении плагиата, а именно: поиск кандидатов, подробное сравнение и последующая обработка на основе знаний. Стандартное обнаружение плагиата рабочий процесс представлен на рис. 2 [6].

На этапе поиска кандидата применяется эвристика, чтобы уменьшить пространство поиска (в том числе тематическая/жанровая фильтрация потенциальных исходных документов).

Potthast et al. (2011) [6] изложил три подхода: первый предполагает формулировку запроса на основе извлеченных ключевых слов из подозрительного документа и переведенного на соответствующий язык; следующие два подхода полагаются на результаты машинного перевода и используют либо стандартный поиск ключевых слов или хэш-кодирование. Подробный этап сравнения включает измерение сходства между подозрительными текстами и потенциальными исходными документами.

Соответствующие методы описаны Potthast et al. (2011) [6]: синтаксический (*CL – CNG*), основанный на словарях (на основе тезауруса на *Eurovoc, CLVSM*), на основе параллельных корпусов (*CL – ASA, CL – LSI, CL – KCCA*) и сопоставимых корпусов (*CLESA*).

Модели на основе синтаксиса

Модель Cross-Language Character N-Gram (*CL-CnG*) основан на работе McNamee и Mayfield (2004) [12]. *CL – CNG* модель использует токенизацию. Размер N-грамм выбирают равным 4. Тексты затем сегментируются в 3-Gram (последовательности из 3 смежных символов) и преобразуются в векторы характера *tf.idf*. Метрика, использу-

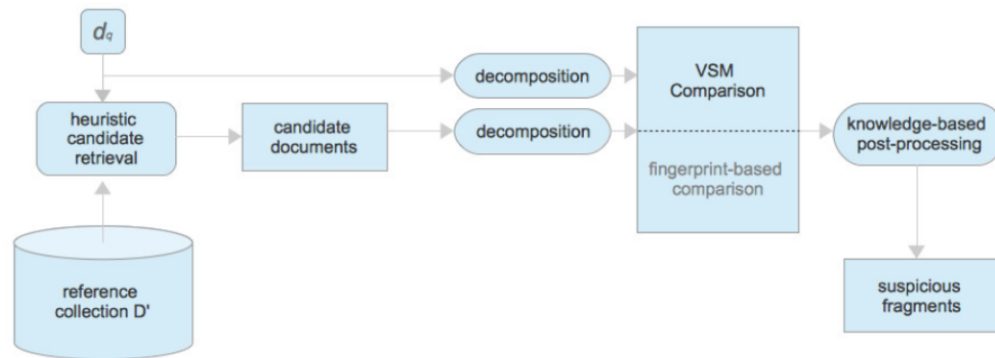


Рис. 2: Процесс обнаружения плагиата (adapted from Potthast et al. (2011)).

емая для сравнения двух векторов, является сходством косинусов. Ключевое отличие этого подхода заключается в возможности сравнения многоязычных документов без перевода. Наилучшие результаты были достигнуты для языков, имеющих сходную синтаксическую структуру и международный лексикон (например, родственные пары европейских языков).

Модели на основе словаря

Модель CL-VSM (Cross-Language Vector Space Model, векторное межъязыковое пространство) заключается в построении моделей векторного пространства документов с использованием индексированных тезаурусов, словарей и другие понятийных пространств. Например, многоязычный JRC-Acquis Parallel Corpus, представленный в Steinberger (2012) [24] связывают тексты с помощью так называемых «языково-независимых анкеров», многоязычных пар слов, которые обозначают имена существительных, местоположения, даты, единицы измерения и т. д.

Торрейон и Рамос (2011) [23] представили комбинацию n-грамм и подхода, основанного на словарях, как расширение системы «CoReMo», разработанной ранее для обнаружения внешнего плагиата. Direct2stem и stem2stem встроены в систему и основаны на Викисловаре и Словаре межъязыковых ссылок Wikipedia. Direct2stem принимает полные слова как записи и предо-

ставляет переводы наиболее часто встречающихся корней в качестве выходных данных. Stem2stem активируется в случае, если предыдущий словарь не смог найти вариант перевода: в этом случае принимаются исходные корни. Если оба словаря терпят неудачу, слово принимается по английским правилам.

Pataki (2012) [11] представляет языковой подход, который состоит из трех основных этапов, а именно: поиск сокращения пространства, оценка подобия и фильтрация результатов. Для сокращения пространства используются средства предварительной обработки документов (фрагментация, устранение стоп-слов), извлечение ключевых слов и перевод их лемм. Цель метода – проверить, могут ли сравниваемые документы быть переводами друг друга.

CL-CTS (Cross-Language Conceptual Thesaurus-Based Similarity) [13] (Гупта, 2012) предлагает метод оценки подобия, который является алгоритмом измерения сходства между текстами, написанными на разных языках (английский, немецкий и испанский языки) на основе доменного отображения представленного в Eurovoc. Функция ад-хок определяет является ли документ частью какого-либо тезауруса понятий id, представленное векторным размером в многомерном векторном пространстве.

CL-CTS нацелен на измерение семантического сходства с использованием абстракт-

ных понятий из слов в текстовых единицах. Метод использует идею Патаки (2012) [11], который для каждого предложения строит мешок слов путем получения всех доступных переводов каждого слова предложения. Мешок слов предложения – это слияние сумм слов предложения. После этого используют расстояние Jaccard (Jaccard, 1912) с нечетким совпадением между двумя мешками для измерения сходства между двумя предложениями.

Модели на основе сопоставимых корпусов

Cross-Language Explicit Semantic Analysis (CL-ESA) основан на явной семантической модели анализа, впервые введенной Габрилович и Маркович (2007) [14], которая представляет значение документа на основе вектора на основе словаря, полученного из Википедии, чтобы найти документ внутри корпуса. Он был повторно использован Potthast и др. (2008) [15] в контексте межъязыкового поиска документов.

CL-ESA представляет собой подход, основанный на сопоставимых корпусах. Согласно Талвенсаари (2008) [16], в отличие от параллельных корпусов (CL-LSI, CL-KCCA и CL-ASA модели), концепция сопоставимых корпусов не связана с переводами, ориентированными на предложения. Она представлена темами, связанными общим лексиконом. Энциклопедия Википедии и аналогичные ресурсы могут служить примером. Эти корпуса являются более шумными, но в то же время более гибкими. CL-ESA подход подразумевает автоматическое создание ассоциаций слов для двуязычного представления документов с целью сравнения лексической корреляции.

Как объясняет Cimiano et al. (2009) [17], пространство концепций C связано с пространством статей в Википедии, поэтому подход называется «явным». Связь между подозрительным документом и концептуальным пространством оценивается путем вычисления суммы величин $tf-idf$ для всех слов анализируемого текста.

Один из последних подходов под названием CL-KGA (Cross-Language Knowledge Graph Analysis) или анализ графов знаний на разных языках, представленный Franco-Salvador et al. (2013) [18], существенно отличается тем, что он основан на использовании недавно созданной многоязычной семантической сети BabelNet и включает графическое представление текста и сравнение. В BabelNet синсеты WordNet и страницы Википедии образуют концепции (узлы), между которыми располагаются семантические указатели и гиперссылки отношений (ребер) соответственно, как объяснено Navigli (2012) [19].

BabelNet [25] — многоязычная сеть большого размера, формируемая с помощью направленных и взвешенных графов, в которых узлы представляют собой понятия и имена объектов, а ребра — отношения между ними. Кроме того, в BabelNet каждый из узлов имеет набор лексикализации понятий в разных языках, давая базу знаний в многоязычном измерении. Список понятий BabelNet формируется всеми значениями слова WordNet и записи Википедии.

Franco-Salvador (2013), используя BabelNet для разработки подхода CL-KGA выяснили, что существует много ложных срабатываний при обнаружении. Авторы предположили, что происходит потеря слов, потому что Словарь BabelNet их не находит, а это, в свою очередь, приводит к тому, что фрагменты документов не обрабатываются. Показано, что использование BabelNet приводит к потере 50% слов, в отличие от статистического словаря обучаемой модели IBM M1, которая находит 70%. Текстовая фрагментация на этапе предварительной обработки проводилась с использованием скользящего окна из 5-предложений, грамматические категории выделяли инструментом TreeTagger. Сходство измерялось на основе значений соотношения и весов концепции.

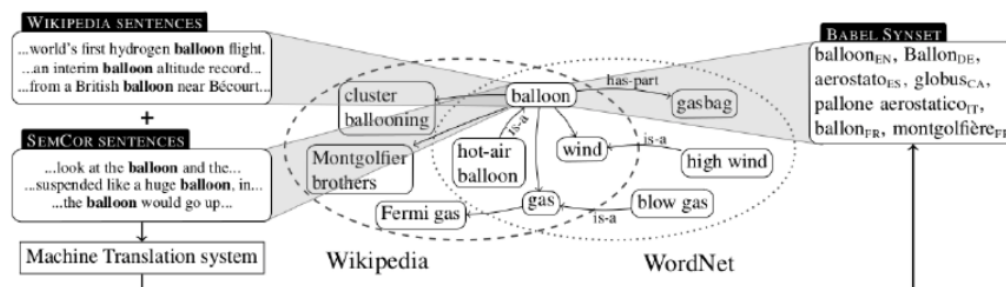


Рис. 3: Пример внутренней структуры BabelNet (рисунок, взятый из [25]).

Модели на основе параллельных корпусов

Cross-Language Alignment-based Similarity Analysis (CL-ASA) был введен Barrón Cedeño et al. (2008 г.) [2] и впоследствии расширен Pinto et al. (2009) [20]. CL-ASA подразумевает создание двуязычного статистического словаря на основе параллельного корпуса с использованием модели IBM 1 (Brown et al., 1993) [21]. Модель предназначена для определения того, является ли текстовая единица переводом другой текстовой единицы с использованием двуязычного словаря, который содержит пары переводов из параллельных корпусов.

CL-LSI (Cross-Language Latent Semantic Indexing) также использует параллельные корпуса. Это общая стратегия, применяемая в IR-системах для ассоциации терминологических документов. Potthast et al. (2011) отмечают, что CL-LSI характеризуется плохой производительностью из-за использования методов линейной алгебры, разложения сингулярного значения оригинала матриц документов, как основы алгоритма [6].

CL-KCCA (Cross-Language Kernel Canonical Correlation Analysis) работает намного лучше, чем LSI на одних и тех же наборах данных, хотя он также основан на SVD, согласно Vinokourov et al. (2002) [22]. CL-KCCA анализирует соответствие точек в двух пространствах вложений, представляющих двуязычные пары документов и измеряет корреляцию соответствующих проекционных значений. Это обеспечивает обнару-

жение определенных смысловых сходств, представленные наборами слов с одинаковыми шаблонами значений для данной двуязычной пары документов.

Модели на основе машинного перевода

Barrón-Cedeño et al. (2013) добавил к классификации пятую модель (MT-based models). Translation + Monolingual Analysis (T+MA) Перевод + Моноязычный анализ (T + MA) состоит в переводе подозрительного плагиата назад, на язык исходного текста, для того, чтобы провести одноязычное сравнение между ними. Muhr et al. (2010) предложили реализацию модели, заключающуюся в замене каждого слова одного текста на его наиболее вероятные переводы на язык другого текста, что приводит к мешку слов. Метрика, используемая для сравнения двух текстов — это одноязычное соответствие, основанное на строгом пересечении мешков слов.

Формальное описание задачи автоматического обнаружения плагиата

Первым из элементов, рассматриваемых при решении задачи автоматического обнаружения плагиата, является текстовый корпус D , который составлен из набора документов. Достижение эталонного корпуса является одним из ключевых факторов успеха систем обнаружения плагиата на основе корпуса. Каждый документ $d \in D$ является потенциальным источником текста, включенного в подозрительный документ, то есть плагиата. Вторым элемент — рас-

смотрение подозрительного документ, который будет известен как s . Этот документ может быть оригинальным, содержать плагиатные фрагменты или, по сути, быть полностью плагиатом.

Цель автоматического обнаружения плагиата — найти документ $d \in D$, который использовался в качестве источника для получения документа s , который предположительно является случаем плагиата. Этот поиск можно вывести на более конкретный уровень: будь $s_i \in s$ фрагментом плагиата, цель состоит в том, чтобы найти такой фрагмент $d_j \in d$, что d_j является источником плагиатированного фрагмента s_i [4].

Справочный набор данных, необходимый для исследования межъязыкового плагиата, должен содержать тексты на разных языках. Так, например, Ферреро и др. предложен набор данных (2016) [5]. Набор данных был специально разработан для обнаружения и оценки межъязыкового текстового сходства. Набор данных многоязычный: он содержит французский, английский и испанские тексты. Набор основан как на параллельных, так и на сопоставимых корпусах (смесь Википедии, статьи научных конференций, обзоры продуктов amazon, Europarl и JRC в разном соотношении). Он содержит тексты как после человеческого, так и машинного перевода, причем документы были переведены разными типами авторов (от среднего уровня до специалистов).

Авторский алгоритм межъязыкового семантического сходства

Следует отметить, что в научной литературе практически отсутствуют систем оценки межъязыкового семантического сходства с русскими текстами и, тем более, межъязыкового плагиата для русских текстов. Одной из таких систем является авторская разработка для сопоставления текстов на русском, украинском и английском языках. Авторами предложен алгоритм и разработана программа обнаружения межъязыкового семантического сходства, относящий-

ся ко второй группе моделей (модели на основе словаря), предварительное название алгоритма — Cross-Language Megalemma-based Similarity Analysis (CL-MSA). Алгоритм основан на предварительном преобразовании текстов в последовательность конструкций (мегалемм), не зависящих от языка.

Идея состоит в том, чтобы преобразовать текст в последовательность конструкций, называемых здесь мегалеммами, которые не зависят от языка, и дальнейшие оценки производить на этих конструкциях.

Мегалемма — это множество морфологически нормализованных однокоренных слов русского языка (derivativная группа) вместе с вариантами перевода этих слов на другие языки. Сейчас другими языками являются украинский и английский. Набор данных для исследований был разработан с применением методов машинного перевода. Derивативные группы терминов были дополнены английскими словами при помощи перевода через yandex-транслятор.

Например, мегалемма ПОДХОД содержит русские слова ПОДХОД, ПОДХОЖДЕНИЕ, ПОДХОДНЫЙ, ПОДХОДИТЬ, ПОДХОДИВШИЙ, ПОДХОДЯЩИЙ, ПОДОЙТИ, ПОДОШЕДШИЙ, украинские аналоги ПІДХІД, ПІДХОДИВ, ПІДХОДИТИ, ПІДХІДНИЙ, ПІДХОДЯЩИЙ и английские APPROACH, GOING UP.

В настоящий момент составлен словарь из около 14000 мегалемм, покрывающий наиболее употребляемые слова.

Мегалемма в общем случае соответствует одному слову исходного текста, однако более информативными являются такие конструкции, как именные группы и генетивные группы. **Именная группа** — это существительное с возможными прилагательными, согласованными по роду, числу и падежу с существительным. Генетивная группа — это две рядом стоящие именные группы, причём вторая в родительном падеже (для английского — разделённая предлогом of).

Для каждой такой группы формируется последовательность мегалемм, которая сортируется внутри группы в алфавитном порядке – эту последовательность назовём мегатокеном. Таким образом, текстам «продуктовый магазин», «магазин продуктов», «магазин продуктів», «shop products» соответствует один и тот же мегатокен [МАГАЗИН + ПРОДУКТ].

Семантическим вектором для текста назовём вектор его мегатокенов с весами. В качестве весов можно выбрать различные метрики, например, частоты встречаемости в этом тексте или tf-idf в некотором корпусе текстов. Для оценки близости текстов можно использовать, например, косинус близости этих векторов. И действительно, для переводных и плагиатных текстов этот коэффициент имеет относительно большое значение.

В качестве коэффициентов вектора терминов были взяты их частоты встречаемости, нормированные на корень из суммы квадратов. Коэффициент близости векторов на одинаковых текстах равен 100, на абсолютно разных – 0.

Семантические вектора могут позволить искать близкие тексты в корпусах довольно больших объёмов, однако для задачи обнаружения плагиата на практике можно использовать следующий подход. Индекс по большой коллекции строится на основе семантических векторов, однако большое значение косинуса ещё не гарантирует наличие плагиата. Верно обратное – наличие плагиата приводит к большому значению косинуса, поэтому можно отобрать некоторое количество «кандидатов» с большим коэффициентом близости для дальнейшей проверки на плагиат более трудоёмким способом на уровне повторов мегалемм. Такой подход использовался в проекте DISSERCOMP (www.dissercomp.ru) в рамках совместной работы с сообществом Диссернет, где он показал свою работоспособность. Как выглядит русско-украинский плагиат можно посмотреть на рис. 4.

На рис. 4 утилита сравнения показывает изменения плагиатного фрагмента по сравнению с исходным фрагментом, но поскольку фрагменты на разных языках, то формально совпадений мало, а на уровне мегалемм тексты практически идентичны.

По предварительным данным, разработанная программа сравнения текстов показала 70% семантического подобия на документах научного содержания и их качественных переводах (исключая именованные сущности).

Заключение

Таким образом, в настоящее время выделяют пять семейств поисковых моделей, основанных на: синтаксисе, тезаурусе, сопоставимых корпусах, параллельных корпусах и машинном переводе.

Результаты оценки эффективности описаны Potthast et al. (2011 г.). CL-C3G, CL-ESA и CL-ASA были сопоставлены для языковых пар: английский как первый язык, испанский, немецкий, французский, голландский или польский как второй [5]. CL-C3G и CL-ESA показывают лучшие результаты, когда выполняется сопоставление информации подозрительных и оригинальных документов. CL-ESA, по сравнению с CL-ASA и CL-C3G, оказался более универсальным методом, однако он во многом зависит от языковых пар. CL-ESA показывает лучшие результаты по сопоставимым корпусам. CL-ASA получает лучшие результаты на параллельных корпусах.

Barrón-Cedeño (2012) сравнивали CL-CNG, CL-ASA и T + MA относительно отдалённых языковых пар (баскско-английский и баскский-испанский). T + MA превосходит другие модели, потому что не зависит ни от корпусов, ни от синтаксического / лексического сходства между языками.

Barrón-Cedeño et al. (2013) представили ещё одну оценку эффективности CL-CNG, CL-ASA и CL-T + MA на корпусе испанский-английский языки. Показано, что производительность моделей зависит от длины до-

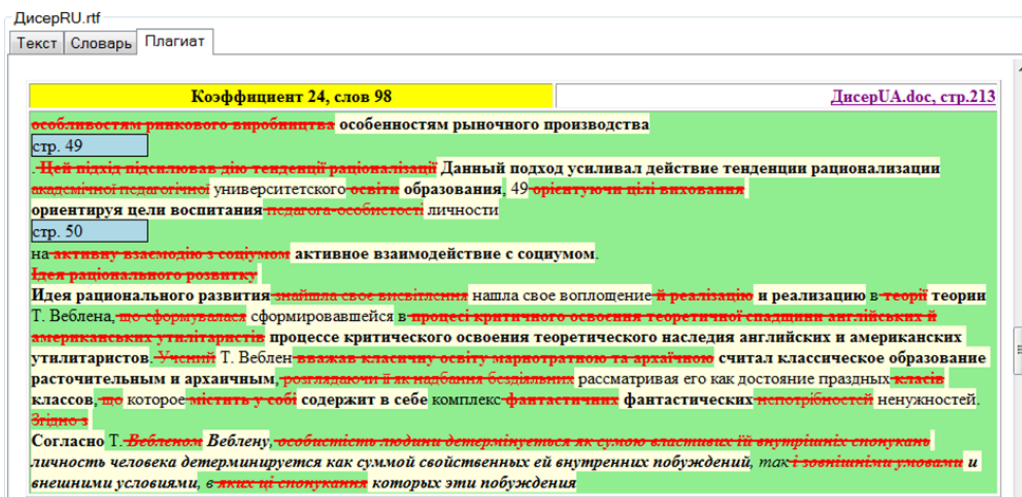


Рис. 4: Результат выявления русско-украинского плагиата

кумента: CL-CNG и T + MA работают лучше с более длинными документами в отличие от CL-ASA.

Детальное исследование Ferrero et al. [5] выявило следующие результаты. CL-ASA, CL-CTS и T + MA являются очень эффективными, но их поведение зависит от уровня гранулярности. Авторы сделали вывод, что изученные методы могут вести себя немного по-разному в зависимости от рассматриваемого текстового блока (фрагмент или предложение), но они ведут себя практически эквивалентно, независимо от языков сравниваемых текстов, если достаточное количество лексических ресурсов доступно для работы с этими языками.

Ferrero et al. показали общее поведение методов CL-ASA, CL-CTS и T + MA на разных языковых парах. Авторы предположили, что если метод эффективен для конкретной языковой пары, он будет таким же эффективным для другой языковой пары при достаточном количестве доступных лексических ресурсов [5].

Potthast et al. предположили, что некоторые методы дополняют друг друга, поэтому комбинация этих методов предполагает очень многообещающие результаты [6]. Ferrero et al. показали, что методы ведут себя по-разному в сопоставлении кластеров и единиц, что открывает новые возможности для их комбинации или слияния [5].

Таким образом, можно заключить, что прототип процесса обнаружения межъязыкового плагиата остается неизменным: он включает эвристическое извлечение, подробное сравнение и фильтрацию, основанная на знаниях. Алгоритмы поиска и сравнения изменяются и базы знаний расширяются.

Торрейон и Рамос (2011) [23] и Патаки (2012) [11] используют специальные словари; Штейнбергер (2012) [24] и Gupta (2012) [13] описывают приложение тезауруса Eurovoc; CL-ESA использует сопоставимые корпуса. Такие модели, как CL-ASA, CL-KCCA, CL-LSI требуют наличия параллельных корпусов для правильного выполнения анализа; CL-KGA подход полагается на использование семантической сети BabelNet, которая объединяет синсеты WordNet с статьями Википедии. MT + A, согласно Barrón-Cedeño et al. (2013) [10], обеспечивает наилучшие результаты, однако весь справочный сбор слишком дорогостоящ, и соответствующие услуги перевода далеко не совершенны.

Предлагаемый нами алгоритм обнаружения межъязыкового семантического сходства относится ко второй группе моделей (модели на основе словаря), предварительное название алгоритма — Cross-Language Megalemma-based Similarity Analysis (CL-MSA). Алгоритм работает для русского, украин-

ского и английского языков. Алгоритм основан на предварительном преобразовании текстов в последовательность конструкций (мегалемм), не зависящих от языка. Последовательность мегалемм, сформированная для именных и генетивных групп, является мегатокеном. Один и тот же мегатокен соответствует разноязыковым текстам. Использование вектора мегатокенов с весами в качестве семантического вектора для текста позволяет оценить семантическую близость текстов на разных языках.

Литература

- [1] Salha Alzahrani, Naomie Salim, Ajith Abraham, and Vasile Palade, “iPlag: Intelligent Plagiarism Reasoner in Scientific Publications”, IEEE World Congress on Information and Communication Technologies, 2011.
- [2] Alberto Barrón-Cedeño, Paolo Rosso, David Pinto, and Alfons Juan. 2008. On Cross-lingual Plagiarism Analysis using a Statistical Model. In Benno Stein and Efstathios Stamatatos and Moshe Koppel, editor, Proceedings of the ECAI’08 PAN Workshop: Uncovering Plagiarism, Authorship and Social Software Misuse. Patras, Greece, pages 9–13.
- [3] Martin Potthast, Matthias Hagen, Anna Beyer, Matthias Busse, Martin Tippmann, Paolo Rosso, and Benno Stein. 2014. Overview of the 6th International Competition on Plagiarism Detection. In PAN at CLEF 2014. Sheffield, UK, pages 845–876.
- [4] Luis Alberto Barrón Cedeño. Detección automática de plagio en texto. Valencia, noviembre de 2008. 76 p.
- [5] Jérémy Ferrero, Frédéric Agnés, Laurent Besacier, and Didier Schwab. 2016. A Multilingual, Multi-style and Multi-granularity Dataset for Cross-language Textual Similarity Detection. In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC’16). European Language Resources Association (ELRA), Portoroz, Slovenia, pages 4162–4169. ISLRN:723-785-513-738-2. <http://islrn.org/resources/723-785-513-738-2/>.
- [6] Martin Potthast, Alberto Barrón-Cedeño, Benno Stein, and Paolo Rosso. 2011. Cross-Language Plagiarism Detection. Language Resources and Evaluation 45(1):45–62. <https://doi.org/10.1007/s10579-009-9114-z>.
- [7] Vera Danilova. Cross-Language Plagiarism Detection Methods. Proceedings of the Student Research Workshop associated with RANLP 2013, pages 51–57, Hissar, Bulgaria, 9-11 September 2013.
- [8] Alberto Barrón-Cedeño. 2012. On the Mono- and Cross-Language Detection of Text Re-Use and Plagiarism (Thesis). Departamento de Sistemas Informáticos y Computación, Universidad Politécnica de Valencia.
- [9] Norman Meuschke, and Bela Gipp 2013. State-of-the-art in detecting academic plagiarism. International Journal for Educational Integrity Vol. 9 No.1, pp. 50-71.
- [10] Alberto Barrón-Cedeño, Parth Gupta, and Paolo Rosso. 2013 Methods for cross-language plagiarism detection. Knowledge-Based Systems 50, 211-217.
- [11] Máté Pataki. 2012. A New Approach for Searching Translated Plagiarism. In Proceedings of the 5th International Plagiarism Conference. Newcastle, UK, pages 49–64.
- [12] Paul McNamee and James Mayfield. 2004. Character N-Gram Tokenization for European Language Text Retrieval. Information Retrieval Proceedings 7(1-2):73–97.
- [13] Parth Gupta, Alberto Barrón-Cedeño, and Paolo Rosso. 2012. Cross-Language High Similarity Search Using a Conceptual Thesaurus. ACLEF 2012, LNCS 7488, pp. 67-75, 2012.
- [14] Evgeniy Gabrilovich and Shaul Markovitch. 2007. Computing Semantic Relatedness using Wikipediabased Explicit Semantic Analysis. In Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI’07). Morgan Kaufmann Publishers Inc., Hyderabad, India, pages 1606–1611.
- [15] Martin Potthast, Benno Stein, and Maik Anderka. 2008. A Wikipedia-Based Multilingual Retrieval Model. In 30th European Conference on IR Research (ECIR’08). Springer, Glasgow, Scotland, volume 4956 of LNCS of Lecture Notes in Computer Science, pages 522–530.
- [16] Tuomas Talvensaari. 2008. Comparable Corpora in Cross-Language Information Retrieval (Academic Dissertation). Acta Electronica Universitatis Tamperensis 779.
- [17] Philipp Cimiano, Antje Schultz, Sergey Sizov, Philipp Sorg, and Steffen Staab 2009. Explicit Versus Latent Concept Models for Cross-Language Information Retrieval. Proceedings of the Twenty-First International Joint Conference on Artificial Intelligence (IJCAI-09).
- [18] Marc Franco-Salvador, Parth Gupta, and Paolo Rosso. 2013. Cross-Language Plagiarism Detection Using a Multilingual Semantic Network. IECIR 2013, LNCS 7814, pp. 710-713.
- [19] Roberto Navigli 2012. Babelplagiarism: What can BabelNet do for cross-language plagiarism detection? Keynotes for PAN 2012: Uncovering, Authorship, and Social Software Misuse.
- [20] David Pinto, Jorge Civera, Alfons Juan, Paolo Rosso, and Alberto Barrón-Cedeño. 2009. A Statistical Approach to Crosslingual Natural Language Tasks. Journal of Algorithms 64(1):51–60. <https://doi.org/10.1016/j.jalgor.2009.02.005>.
- [21] Peter F. Brown, Vincent J. Della Pietra, Stephen A. Della Pietra, and Robert L. Mercer. 1993. The Mathematics of Statistical Machine Translation: Parameter Estimation. Computational Linguistics 19(2):263–311. <http://www.aclweb.org/anthology/J93-2003>.

- [22] Alexei Vinokourov, John Shawe-Taylor, and Nello Cristianini. 2002. Inferring a Semantic Representation of Text via Cross-Language Correlation Analysis. *Advances of Neural Information Processing Systems* 15.
- [23] Diego Antonio Rodríguez Torrejón, and José Manuel Martí Ramos. 2011. Crosslingual CoReMo System. Notebook for PAN at CLEF 2011.
- [24] Ralf Steinberger 2012. Cross-lingual similarity calculation for plagiarism detection and more – Tools and resources. Keynotes for PAN 2012: Uncovering, Authorship, and Social Software Misuse.
- [25] Navigli R. & Ponzetto S. P. (2010). Babelnet: building a very large multilingual semantic network. In proceedings of the 48th annual meeting of the Association for Computational Linguistics, ACL '10. Stroudsburg, PA, USA, 216–225.

Применение LSA/LSI методов для извлечения данных из неструктурированного текста*

Е. Г. Соколов¹, М. М. Шарнин²

evgeny.sokolov@phystech.edu¹, mc@keywen.com²

¹Московский физико-технический институт (государственный университет).

²Институт проблем информатики Федерального исследовательского центра "Информатика и управление" Российской академии наук.

В работе приводятся результаты применения моделей семантического анализа для задач поиска близких по смыслу фрагментов в неструктурированном тексте. Также рассматриваются возможные варианты улучшения работы алгоритмов.

Ключевые слова: Семантика, NLP, LSA, LSI, токен, факторный анализ, SVD, гипотеза компактности.

Application of LSA/LSI models for data extracting from raw text*

E. G. Sokolov¹, M. M. Sharnin²

¹Moscow Institute of Physics and Technology (State University)

²Institute of Informatics Problems FRC CSC of the Russian Academy of Sciences. Moscow. Russia

The article describes practical aspects of Latent Semantic Analysis (LSA) methods for search and extracting semantic structures from raw text. Also considered are possible options for improving the work of algorithms.

Key words: Semantic, NLP, LSA, LSI, token, factor analysis, SVD, compact hypothesis.

1. Введение

Автоматический анализ информации в неструктурированном тексте (NLP) - проблема, возникающая во всех областях, работающих с большими объёмами данных, написанными на одном из естественных языков (правительственные структуры, социальные институты, образование и т.д.). Одной из самых распространённых задач является обнаружение наличия/отсутствия определённого семантического смысла в большом объёме текстовых данных без непосредственного прочтения человеком.

Для решения задач такого рода могут использоваться следующие подходы: 1) построение полной дистрибутивно-семантической модели [1]; 2) LSA - латентно-семантический анализ и LSI - латентно-семантическое индексирование (название LSA-подхода в области информационного поиска).

В сравнении с первым подходом, LSA позволяет строить семантические зависимости, основываясь на меньшем корпусе тек-

стов, которые также могут состоять из большого количества специфических терминов и определений, редко встречающихся в неспециализированных информационных источниках. Таким образом, LSA позволяет построить семантические зависимости на тех объёмах данных, при которых построение полной дистрибутивной семантической модели невозможно. Хотя на достаточных объёмах данных LSA проигрывает в точности.

Оригинальный алгоритм был предложен в 1990-м году группой американских исследователей [2]. Впоследствии данный подход был распространён на широкий класс задач.

Приведём некоторые примеры успешного использования метода для схожих задач.

1. Использование LSI для построения графов террористических организаций [3]. В данном исследовании представлен автоматизированный подход для генерации и отображения оценки узлов и ссылок, относящихся к выбранной теме. Эта работа сочетает извлечение именованных объектов и LSI-подход.

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 16-07-00756, 16-29-09527, 18-07-00909, 18-07-01111, 18-07-20030, 18-37-10008.

2. Визуализация авторства документов при помощи N-граммов и LSI [4]. С помощью LSI документы и термины представляются в векторном пространстве низкой размерности и затем могут быть визуализированы.
3. Автоматическая аннотация изображений при помощи LSA-модели [5]. Использование классической модели LSA для аннотации изображений и сравнение точности с pLSA (probabilistic LSA).
4. Использование LSA для идентификации схожих фрагментов исходного кода [6]. В данной работе LSA используется для извлечения и сравнения исходного программного кода с описанием из документаций.
5. Использование LSI в качестве спам-фильтра [7]. Приводятся результаты экспериментов, которые оправдывают применимость подхода для этой задачи.
6. Межъязыковой патентный поиск [8]. В исследовании приводятся результаты применения LSA/LSI-подходов, адаптированных для группировки ключевых терминов и поиска по ним в мультиязычном пространстве документов.

В данной же работе проверяется гипотеза, что метод применим для поиска фрагментов текста с близким смыслом в производственных отчётах, несмотря на различие терминологии.

2. LSA

Латентно-семантический анализ - метод обработки текстовой информации, способный выявлять семантические связи между корпусами текстов и терминами, в них встречающимися. Основан данный метод на факторном анализе, теореме о сингулярном разложении матриц (SVD-разложении) [9] и гипотезе компактности.

В общем виде, согласно теореме о сингулярном разложении, любая вещественная прямоугольная матрица может быть разложена на произведение трёх матриц [9]. В дан-

ном случае, взяв X - матрицу, описывающую частоту терминов, употребляющихся в корпусе текстов, получим:

$$X = T_0 S_0 D_0^T,$$

где T_0 и D_0^T - ортогональные матрицы, а S_0 - диагональная матрица.

У такого SVD-разложения есть свойство, что если в исходной матрице S_0 оставить только n наибольших сингулярных значений и в матрицах T_0 и D_0^T соответствующие столбцы, то наилучшим приближением матрицы X к матрице \hat{X} ранга n будет:

$$X \approx \hat{X} = T S D^T,$$

где S - матрица с n наибольшими сингулярными значениями, T и D^T - матрицы с соответствующими значениями столбцов.

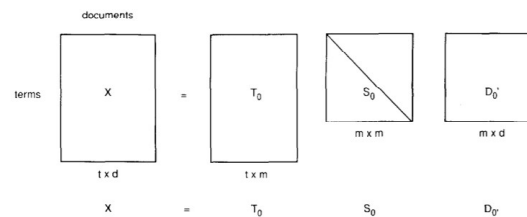


Рис. 1: SVD-разложение терм-документной матрицы [2].

Таким образом, термины и документы оказываются представленными в векторном пространстве размерности n .

Выбор оптимального значения n до сих пор остаётся открытой проблемой, и в каждой конкретной задаче подбирается эмпирически, исходя из лучшего полученного результата. Более подробно о работе LSA и подборе n написано в работах [10,11].

3. Описание задачи

В качестве текстовых данных для анализа в работе используются ежедневные отчёты с буровых установок. В них инженеры в свободной форме, с использованием узкотехнических терминов, описывают события, которые происходили в их смену. Далее эти отчёты направляются аналитикам для их изучения. Необходимо по наличию/отсутствию инцидентов, описанных

в тексте, скорректировать планы по разработке месторождений. Как уже было сказано, описание самих аварий и их наличие/отсутствие описывается людьми в свободной форме на английском языке, но наличие системы для анализа и поиска информации по инцидентам в исходных данных может сильно облегчить и ускорить работу отделов аналитики и планирования.

Сам отчёт в конечном виде представляет собой PDF-документ, основной частью которого является таблица, в столбцах которой находятся такие данные, как время начала работ, продолжительность, глубина и ещё ряд параметров, характеризующих процесс добычи. Последней колонкой является описание процессов, характеризующих процесс бурения в данный промежуток времени. Описание выполняется в свободной неструктурированной форме. Данная колонка и будет конечной целью для анализа методом LSA.

Start	Hours	Depth	Description
t_s	t				d			Data

Рис. 2: Структура таблиц в отчётах, где t_s - время начала события/цепочки событий, t - продолжительность, d - глубина, Data - текстовое описание событий.

4. Решение с использованием LSA

В данной работе решение задачи написано на языке Python 3 с использованием библиотеки Gensim, которая представляет собой ряд инструментов, упрощающих работу с большими массивами текстовых данных.

В начале, распарсив PDF-документы, получаем текстовые описания, содержащиеся в последней колонке таблицы (рис.2). Это та информация, на которой будет строиться семантическая модель посредством LSA. Далее нам необходимо разбить весь текст на отдельные последовательности символов (токены) и убрать наиболее часто встречающиеся союзы, местоимения, предлоги и т.д., дабы они в дальнейшем не участвовали в формировании семантических связей. Также необходимо удалить те токены, кото-

рые встречаются всего лишь 1 раз в корпусе текстов, т.к. полезной информации они не могут принести из-за редкости употребления, но могут добавить лишний шум.

Далее необходимо представить текст, разбитый на абзацы, в векторном виде. Для этого воспользуемся методом, который называется «мешок слов» (bag-of-words). В конечном итоге, данные будут представлены в виде матрицы соответствий предложений и терминов. Таким образом, каждое предложение будет описано n -мерным вектором, где число n будет зависеть от финального количества токенов.

Наконец, конвертируем предложения и интересующие нас термины (в данном случае это типы аварий на буровых) в n -мерное пространство. В дополнение, введём функцию оценки близости запросов и предложений с помощью косинусной меры близости.

5. Результаты

Пример работы алгоритма на одном из отчётов для поиска 2-х аварийных событий (Mud Losses и Lost Circulation):

1. "Ran in hole with 14.75 Power Drive BHA from 251 m to 1029 m. Filled every 10 stands. No restriction when passing top of 16 liner. at 400 m: P/U 603 kN; S/O 607 kN; ROT 607 kN. at 700m: P/U 691 kN; S/O 660 kN; ROT 673 kN. at 1020 m: P/U 769 kN; S/O 672 kN; ROT 716 kN. Note. Held kick drill. Well secured in 64 sec. No losses."
2. "Broke circulation with 2500 lit/min, 89 bar SPP. Washed down from 1029 m to 1046 m. 40 kN restriction at 1046 m. Picked up, washed down and passed restriction with max 31 kN WOB. Continued to wash down from 1046 m to 1084 m. Note. From 1073 m continuous gradual WOB increase from 4 kN to 50 kN at 1084 m."
3. "Reamed down 1084 m to 1086 m with 3000 lit/min, 116 bar SPP, 60-70 RPM, 8-10 kN-m torque, 9-35 kN WOB."
4. "Drilled 14 3/4 hole from 1086 m to 1118 m (bit depth) with 3000 lit/min, 116 bar

- SPP, 70-100 RPM, 7.7-11.8 kN-m torque, 13-44 kN WOB."
5. "Circulated at 3000 lit/min, 118 bar SPP. Reciprocated string from 1118 m to 1113 m while circulating."
 6. "Observed gradual circulating pressure reduction from 178 bar to 173 bar. Investigated problem. Mud pump B performing at low efficiency. Switched to mud pump C. Unable to pump with mud pump C at required rate. Observed very low pumping pressure. 4-6 bar circulating pressure fluctuation. Raised permitry, performed required isolations. Piece of wood (110x30x15 mm) found in mud pump C suction module valve. Removed same. Meanwhile pumped out of hole from 1302.8 m to 1141 m with 2000 li/min, 39 bar SPP. No over pulls."
 7. "RIH with no rotation/circulation from 1141 m to 1168 m. Gradual increase of WOB up to 44 kN. Broke circulation. Reamed down from 1168 m to 1302.8 m with 30 RPM, 2000 lit/min, 40 bar SPP. No restriction. No torque fluctuation/"

Mud Losses: (6, 0.9305616), (5, 0.85309166), (4, 0.84738356), (3, 0.79817563), (7, 0.67961329), (2, 0.39225531), (1, -0.40293679)

Lost Circulation: (2, 0.92467004), (1, 0.91013885), (7, 0.74197721), (3, 0.61231804), (4, 0.541493), (5, 0.53234458), (6, 0.37769562)

Для каждого термина (x, y): x - номер абзаца с описанием, а y - косинусная мера близости:

$$similarity = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} =$$

$$= \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}},$$

где A и B - два вектора, для которых рассчитывается расстояние.

Таким образом, была построена модель, обученная на отчётах за 14 дней, которая

выдаёт верные предсказания, но их точность недостаточно высока из-за небольшого объёма обучающей выборки.

В приведённом выше примере можно заметить, что для аварийного события (Mud Losses) наиболее вероятными абзацами являются 6, 5 и 4. В 6-м абзаце предложение «Mud pump B performing at low efficiency» напрямую говорит об этом. В абзацах 5 и 4 нет чёткой семантической связи с данным событием, но они описывают те события, которые предшествовали аварийной ситуации и, что наиболее вероятно, имеют схожий семантический характер с подобными цепочками событий, которые характеризовали данный тип аварий в обучающей выборке отчётов. Для аварийного события (Lost Circulation) характерными абзацами являются 2, 1 и 7. Во 2-м и 7-м абзацах непосредственно есть строчка, содержащая «Broke Circulation», которая является семантически очень близкой к искомой. 1-й абзац попал с высокой оценкой, вероятнее всего, по той же причине, что и в случае с Mud Losses.

6. Выводы

Построенная модель подтверждает гипотезу применимости LSA-подхода к решению задачи поиска фрагментов текста с близким смыслом в производственных отчётах. В дальнейшем планируется усовершенствовать модель, переобучив её на больших объёмах данных, а также сравнить результаты с некоторыми модификациями оригинального метода: MASHA (полиномиальный асимметричный иерархический анализ), HLPSA (иерархический вероятностный латентно-семантический анализ).

Литература

- [1] Е.Г. Соколов, А.С. Казанцев, М.М. Шарнин «Применение технологии нейронных сетей и векторных операций для выявления объектов различных классов и их связей в большом корпусе текстов» международная конференция «Ситуационные центры и информационно-аналитические системы класса 4i для задач мониторинга и безопасности» (SCVRT2017), Царьград, 28 ноября 2017 г.

- [2] Scott Deerwester, Susan T. Dumais, George W. Furnas, Thomas K. Landauer, Richard Harshman «Indexing by Latent Semantic Analysis» (1990).
- [3] Bradford, R., Application of Latent Semantic Indexing in Generating Graphs of Terrorists Networks, in: Proceedings, IEEE International Conference on Intelligence and Security Informatics, ISI 2006, San Diego, CA, USA, May 23-24, 2006, Springer, LNCS vol. 3975, pp. 674-675.
- [4] Soboroff, I., et al, Visualizing Document Authorship Using N-grams and Latent Semantic Indexing, Workshop on New Paradigms in Information Visualization and Manipulation, 1997, pp. 43-48.
- [5] Monay, F., and Gatica-Perez, D., On Image Auto-annotation with Latent Space Models, Proceedings of the 11th ACM international conference on Multimedia, Berkley, CA, 2003, pp. 275-278.
- [6] Maletic, J.; Marcus, A. (November 13-15, 2000). "Using Latent Semantic Analysis to Identify Similarities in Source Code to Support Program Understanding". Proceedings of 12th IEEE International Conference on Tools with Artificial Intelligence. Vancouver, British Columbia: 46-53.
- [7] Gee, K., Using Latent Semantic Indexing to Filter Spam, in: Proceedings, 2003 ACM Symposium on Applied Computing, Melbourne, Florida, pp. 460-464.
- [8] Yen-Liang Chen, Yu-Ting Chiu; Cross-language patent matching via an international patent classification-based concept bridge, Journal of Information Science, 2013, pp. 1-17
- [9] Голуб Дж., Ван Лоун Ч. Матричные вычисления. М: «Мир», 1999.
- [10] Thomas Landauer, Peter W. Foltz, Darrell Laham «Introduction to Latent Semantic Analysis» (1998).
- [11] Thomas Landauer, Susan T. Dumais. «A Solution to Plato's Problem: The Latent Semantic Analysis Theory of Acquisition, Induction, and Representation of Knowledge» (1997).

Разработка методов статистического анализа данных для оценки характеристик «визуального восприятия - действия» в системе виртуального окружения тренажеров*

Д. В. Кудрявцев, В. П. Алешин
mitya-kudr@mail.ru

Московский физико-технический институт (государственный университет)

В статье рассматривается одна из гипотез, объясняющая то, как в мозгу происходит обработка изображения, полученного от глаза. Также есть обзор того, как происходят движения глаз и зачем они нужны. Затронута тема обработки изображения мозгом и его стабилизация.

Ключевые слова: обработка изображений, глаз, мозг, зрение

Development of methods for statistical data analysis to assess the characteristics of "visual perception - actions" in the virtual environment of simulators*

D. V. Kudryavtsev, V. P. Aleshin

Moscow Institute of Physics and Technology (State University)

This paper considers one of the hypotheses, explaining how the brain processes the image obtained from the eye. There is also an overview of how eye movements occur and why they are needed. As well, the theme of image processing by the brain and its stabilization is touched.

Key words: Image processing, eye, brain, vision

1. Введение

При создании визуально силовых систем огромное значение играет восприятие информации человеческим глазом. Зрительная система является наиболее важной сенсорной системой человека. Это объясняется тем фактом, что при обработке визуальной информации участвует примерно половина коры головного мозга.

Существует огромное количество работ и гипотез, посвящённых этой теме, но до сих пор до конца не понятно, каким образом наш мозг обрабатывает получаемое изображение. Существует много исследований посвящённых этому вопросу. Нобелевские премии в области медицины и физиологии на тему глаза и мозга получило порядка пяти групп исследователей. Более того, сейчас всё больше и больше развиваются технологии виртуальной реальности, а вместе с ними и 3D-моделирование, стремящееся показывать нам именно то изображение, ко-

торое нужно мозгу для полноценного ощущения реальности смоделированного мира.

Важно понимание того, как статическая (прямая картинка того, что видит глаз) и динамическая (связанная с движением глаз) информация, передаваемая из глаза в мозг, обрабатывается и принимается последним. Это необходимо для объяснения зрительного процесса и создания моделей, которые мы сможем использовать при построении более качественных систем, передающих изображение человеческому глазу.

Обработку изображения глазом и мозгом можно разделить на несколько этапов:

- 1) получение статичного изображения глазом, его отправка в мозг;
- 2) движение глаз, позволяющее удерживать объекты восприятия во времени и позволяющее улучшить статичное изображение;
- 3) работа мозга по стабилизации изображения, полученного от глаза;
- 4) превращение этих изображений в полноценную трёхмерную сцену.

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 16-07-00637, 18-07-20030, 18-37-10008.

Первый этап был частично исследован нами в прошлом учебном году, второй и тре-

тий этапы были изучены моими коллегами. Остаётся изучить четвёртый этап, исправить недостатки исследований первых трёх этапов и составить из них единую картину работы нашего мозга по обработке изображения, полученного из глаз.

2. Зрительная система

Глаз человека по форме напоминает сферу. Наружный слой глазного яблока, называемый склера, представляет из себя твёрдую оболочку, переходящую в передней части в прозрачную оболочку – роговицу. Диптрический аппарат глаза, участвующий в построении изображения на его внутренней поверхности, состоит из роговицы, хрусталика, прозрачной водянистой жидкости и прозрачного стекловидного тела, заполняющих глазное яблоко. К этому аппарату также относится ресничное тело, обеспечивающее изменение кривизны хрусталика, роговица и радужная оболочка, изменяющая размер зрачка. Этот механизм позволяет настраивать резкость и светосилу оптической системы [8]. В итоге изображение попадает на сетчатку, которая и преобразует световую информацию в сигнал, поступающий в мозг через зрительный нерв.

3. Строение глаза

В человеческом глазу есть два типа фоторецепторов, от которых зависит, как мы видим – колбочки и палочки.

Колбочки в основном содержатся в центральной ямке. Они намного чувствительнее к мелким деталям и быстрому движению, нежели палочки, и поэтому отвечают за остроту зрения. Также они отвечают и за цветное зрение.

Палочки, в свою очередь, с большей плотностью расположены на периферии сетчатки, нежели в её середине. Они не различают цвет, однако приблизительно в 100 раз более чувствительны к свету и отвечают за ночное зрение [1]. К слову, поэтому ночью слабый свет легче заметить, если смотреть на него не прямо, а с боку (чтобы свет падал на палочки, а не на колбочки).

Из вышесказанного можно сделать вывод, что изображение, получаемое глазом в первую очередь, по бокам нечёткое и бесцветное, а в центре чёткое и цветное.

Однако это ещё не все недостатки человеческого глаза. Чтобы передать зрительные раздражения от колбочек и палочек в головной мозг нужны нервы. Зрительный нерв находится в каждом глазу и отвечает именно за это. Выходит он из диска зрительного нерва, под углом 15 градусов, считая от центра сетчатки [3]. Однако, поскольку на этом участке сетчатки нет никаких фоторецепторов, при попадании на него света не возникает никакой зрительной реакции. В этом месте возникает так называемое слепое пятно.

Диаметр слепого пятна около 2 мм, что соответствует полю зрения 6 градусов. Это значит, что человек с расстояния 1 м может не увидеть предмета диаметром 10 см, если его изображение проектируется на слепое пятно. Обычно человек этого не замечает, так как мозг сглаживает изображение и слепое пятно заполняется изображением, получаемым движением данного глаза, или изображением от другого глаза.

На сетчатке, кроме слепого пятна, есть ещё одна «аномальная» область: жёлтое пятно, или макула. Она представляет из себя эллипс с радиусом 2,5-3 мм вокруг центральной ямки и несколько выше места входа в глаз зрительного нерва [1]. Эта область имеет жёлтую пигментацию, служащую для защиты сетчатки глаза от ультрафиолета, ведь порядка 80 процентов света, попадающего в глаз, оказывается именно в этой области, отсюда и название.

Плотность фоторецепторов в этой области высокая, хоть и меньше, чем в центральной ямке. В дополнение ко всему вышесказанному свет в человеческом глазу проходит три отдельных преломления. Первое преломление происходит, когда свет пересекает роговицу. Далее свет проходит через хрусталик, имеющий две выпуклые поверхности.

Изображение, проходя сквозь переднюю поверхность хрусталика, снова приводится в порядок. Но потом оно дальше преломляется, поскольку проходит сквозь заднюю поверхность хрусталика, в результате чего возникает конечное изображение на сетчатке, которое является перевернутым вверх дном [2]. В добавок ко всему, изображение на сетчатке получается сферическим из-за строения нашего глаза и положения сетчатки [4].



Рис. 1: Получаемое глазом изображение из внешнего мира (зелёные линии, с разницей в угле в 15 градусов, глаз, конечно, не видит, они нужны для того, чтобы заметить, где находится слепое пятно)

Итого, на рис.1 можно увидеть, какое примерно изображение получает наш глаз (рассмотрим фото МФТИ). И это только для одного глаза.

Кроме всего вышесказанного, у человека может быть ещё много различных заболеваний и дефектов глаз, напрямую влияющих на изображение, получающееся в мозгу: астигматизм, миопия, глаукома, катаракта, конъюнктивит и ещё несколько десятков других. Однако в данном случае, будем считать, что человек имеет идеальное зрение.

Благодаря всему вышесказанному можно понять, как примерно формируется изображение на сетчатке, то есть какое примерно изображение передаётся в мозг. Это даёт понимание первого этапа обработки изображения глазом и мозгом, который был представлен во введении.

4. Движение глаз

В процессе движения глаз принимает участие шесть мышц, 2 косые и 4 прямые — наружная, внутренняя, верхняя и нижняя. Каждая из них играет определенную роль в регулировании положения глазного яблока.

Глаз человека может перемещаться вокруг любой оси, проходящей через центр вращения, который расположен примерно в 13,5 мм от верхушки роговицы [8]. Движение глаз совершается всегда сообща, т.е. зрительные оси левого и правого глаза всегда направлены на один и тот же предмет. При фокусировке на предметах, находящихся на одном расстоянии от глаз, оба глаза совершают одновременные движения.

Существует несколько основных типов движения глаз. Саккады — быстрые движения, которые резко меняют точку фиксации. Саккады могут быть вызваны добровольно, но возникают рефлекторно, когда глаза открыты, даже если они зафиксированы на объекте.

При данном виде движения глазу требуется около 200 миллисекунд для смены положения. Во время этой задержки вычисляется положение объекта относительно центральной, а также разница между исходным и конечным положением преобразуется в команду мышцам, чтобы переместить глаза на правильное расстояние. Саккадические движения глаз считаются баллистическими, потому что система не может реагировать на последующие изменения положения объекта в течение движения глаз [8].

Когда мы смотрим на объект, нам кажется, что наши глаза не совершают перемещений. На самом деле, они постоянно движутся [9]. Существуют различные типы миниатюрных движений глаз, в том числе: щелчки, заносы, нерегулярные движения и высокочастотные толчки. Когда мы фиксируемся на объекте, его образ падает на центральную ямку. Эффект таких небольших движений заключается в том, чтобы постоянно перемещать это изображение по ямке так, чтобы она постоянно стимулировалась.

Движения преследования - сопряженные движения глаз, которые плавно отслеживают медленно движущиеся объекты в поле зрения. Для их совершения требуется движущийся объект, и обычно они не находятся под контролем человека. Также их цель состоит в том, чтобы стабилизировать движущиеся объекты на сетчатке, тем самым позволяя нам воспринимать объект в деталях [8].

Вергентные движения глаз - движения, в которых оба глаза движутся в противоположных горизонтальных направлениях, что позволяет получить изображение ближнего или дальнего объекта. Нистагмы — произвольные ритмичные движения глазных яблок высокой частоты. К нистагмам относят такие микродвижения глаз как тремор и дрейф.

Долгое время считалось, что основная функция движения глаз сводится к тому, чтобы удерживать объект восприятия в поле зрения и менять точки фиксации, расширяя общий угол обзора. Однако исследования [9] показали, что объекты восприятия, которые оказываются неподвижными относительно сетчатки, перестают восприниматься фоторецепторами. Это обуславливается тем, что импульсы в зрительном нерве человека возникают лишь в ответ на изменение света, падающего на сетчатку.

5. Отслеживание движения глаза

За последние 50 лет учёные, исследующие движения глаз, добились огромного успеха. К настоящему времени довольно хорошо изучены основные нейронные пути, контролирующие различные типы движений глаз, значительно улучшилась технология отслеживания положения взгляда и, что наиболее важно, стала широко доступной.

Исследования движения глаз получили широкое распространение в различных дисциплинах, от биологии и медицины до информатики и экономики.

Научные исследования, посвящённые движению глаз, начались в конце XIX ве-

ка, когда были впервые разработаны надежные методы измерения положения глаз. Несмотря на то, что некоторые из этих устройств имели замечательную точность измерения, они, как правило, изготавливались по индивидуальному заказу и не были широко доступны.

По записям траекторий Ярбус установил, что скорость дрейфа хаотически меняет свою величину в пределах от нуля до примерно 30 угловых минут в секунду. Во время фиксации, только в результате дрейфа, ось глаза перемещается со средней скоростью, равной приблизительно 6 угловым минутам в секунду и проходит в одну секунду путь, равный 10—15 диаметрам колбочек центральной части fovea.

Важным фактом является то, что несколько раз в секунду на время, длящееся сотые и десятые доли секунды, дрейф достигает почти максимальных величин, т. е. около 30 угловых минут в секунду.

Один из самых популярных методов отслеживания движения глаза, известный как Scleral Coil Tracking [13], был разработан Дэвидом Робинсоном. Данный способ позволяет довольно точно измерять положение глаз и до сих пор используется почти во всех исследованиях, посвящённых физиологии движения глазного яблока.

Суть этого метода состоит в том, что на склере глаза, движение которого мы хотим изучить, накладывается силиконовое кольцо, в которое встроена проволочная катушка, также её называют склеральной. От неё отходит тонкий провод, который соединяется с внешним измерительным блоком, необходимым для измерения напряжения в проволочной катушке. Внешние генераторные катушки создают магнитное поле, которое индуцирует напряжение в проволочной катушке. Отслеживание склеральной катушки зависит от переменных магнитных полей, присутствующих в глазах пользователя.

Использование переменного тока для возбуждения электромагнита, образованного

проволочной катушкой, является эффективным методом генерирования магнитного поля, которое колеблется с определенной частотой. Согласно закону Био-Савара электрический ток, протекающий вдоль проволочной катушки, будет генерировать магнитное поле.

Колебательный магнитный поток от этих генераторных катушек индуцирует ток той же частоты внутри склеральной катушки в соответствии с законом индукции Фарадея. Ток, индуцированный в склеральной катушке, пропорционален скорости изменения магнитного потока через поверхность, ограниченную ею. Поток через катушку зависит от её ориентации в магнитном поле. Если она совмещена с полем (то есть нормальный вектор к катушке совмещен с полем), то поток и величина индуцированного напряжения будут максимальными.

Используя это явление, трекеры оценивают ориентацию глаз, измеряя величину сигнала напряжения в склеральной катушке. Однако, поскольку информации из одного магнитного поля недостаточно для полного определения взгляда, большинство трекеров используют три ортогональные генераторные катушки, работающие на разных частотах. Рассматривая сигнал склеральной катушки на разных частотах, можно оценить магнитный поток, обусловленный каждой из катушек генератора.

6. Модель движения глаз

В последние годы появилось достаточно много исследований, посвящённых построению модели движения глаз для конкретных задач [14]. Самые популярные из них: движение глаз при чтении, при восприятии сцены, при визуальном поиске. У каждой из подзадач есть свои особенности.

Больше всего изучен характер движения глаз при чтении. И для этого есть своя причина. При чтении у зрителя есть определенная задача: люди обычно читают, чтобы понять текст. Это включает в себя последовательность движений глаз, которые обычно

перемещаются слева направо по странице, а затем вниз по странице.

Конечно, задача может быть несколько изменена, например, читателя могут попросить просканировать текст, что приведет к изменению характеристик движения глаза. Тем не менее, большая часть исследований использует понимание как цель читателя.

Когда мы читаем, смотрим на сцену или ищем цель в визуальном массиве, мы перемещаем глаза каждые 250-350 мс. Движения глаз важны для людей, потому что, как я писал выше, высокая острота зрения ограничена центральной ямкой - маленькой круговой областью на сетчатке, которая плотно заполнена фоторецепторами. Движения глаз могут направлять ямку к новым объектам, представляющим интерес.

Человеку кажется, что он может обработать весь визуальный массив информации за одну фиксацию сцены, и хотя он может сделать это достаточно быстро, в действительности мы не сможем полностью обработать информацию за пределами фовеального видения без способности перемещать глаза.

Важным моментом при движении глаз является то, что они являются более или менее баллистическими. После инициирования трудно изменить свою траекторию. Кроме того, поскольку они являются двигательными движениями, требуется время для планирования и выполнения саккады. Исследования показывают, что характер задачи влияет на средний объем времени, затрачиваемого на каждую фиксацию, и среднее расстояние, на которое двигаются глаза. Также для каждой из задач существует значительная вариабельность как в отношении продолжительности фиксации, так и длины саккады.

Существует много различий между чтением, восприятием сцены и визуальным поиском, но также существуют и некоторые факты, которые могут помочь в создании общей модели.

Во-первых, количество информации, которое обрабатывается в любой фиксации, зависит от задачи. При чтении мозгу требуется обработать меньше информации, чем при восприятии сцены или визуальном поиске. Таким образом, например, фиксации при восприятии сцены достаточно длинные, но также достаточно длинные саккады, так как обрабатывается больше информации при одной фиксации.

Во-вторых, трудность восприятия влияет на движения глаз: при чтении, когда текст усложняется, фиксация глаз становится длиннее, а саккады становятся короче; аналогично при восприятии сцены и визуальном поиске. В-третьих, сложность конкретной задачи явно влияет на движения глаз. Наконец, во всех трех задачах есть некоторые доказательства того, что наблюдатель плохо воспринимает информацию во время саккад, но эффективно обрабатывает данные при каждой фиксации.

Недавние исследования показали, что движения глаз при фиксации можно довольно точно описать с помощью Винеровского процесса [10]. Этот алгоритм строится на том факте, что фиксирующие движения глаз сохраняются на коротком промежутке времени [11]. Это значит, что глаз проявляет тенденцию продолжать движение в своем текущем направлении. Поэтому, для реализации модели фиксирующих движений глаз, представляется перспективным реализовать случайное блуждание, которое позволяет избежать перемещения в районы пространства, которые были посещены незадолго до этого.

Из этого и предыдущих разделов, связанных с движениями глаза, можно сделать вывод о том, как движения глаз позволяют удерживать объекты восприятия во времени, это и есть базовый теоретический материал для понимания второго этапа обработки изображения глазом и мозгом, представленного во введении.

7. Обработка изображения мозгом

Поняв, какое изображение на вход получает мозг, стоит понять то, как человеческий мозг обрабатывает его и получает чёткое, цветное и объёмное изображение.

Существует несколько гипотез того, как это происходит. Одна из них – информационный подход, основанный на идеях Дэвида Марра (1945-1980). Компьютеры могут применять похожий подход для распознавания изображений, поэтому представляется возможным создать некоторую математическую модель для него.

Информационный подход основывается на дроблении задачи распознавания объектов на ряд последовательно выполняемых стадий, или операций. И как на конвейере, выход предыдущей операции подаётся на вход последующей. Выделим три этапа: первоначальный эскиз, 2,5-мерный эскиз и представление трехмерной модели [1].

Первоначальный эскиз извлекает общую и абстрактную информацию, отделяет фон и объекты, определяет края этих объектов по прерывистости светимости и интенсивности отражаемого ими света.

Следующий этап определяет ориентацию, глубину контуров, структуру объектов относительно наблюдателя и расстояние до них.

Заключительный этап трансформирует информацию, полученную из предыдущего этапа, в трёхмерную модель, которая преимущественно не зависит от местоположения наблюдателя и его ориентации в пространстве.

Возможный нейронный механизм для исключения шумов, вызванных движением глаз, был обнаружен в сетчатке [12], что может быть одной из причин, почему объекты кажутся стабильными.

Тем не менее, факт того, что детали объекта остаются четкими, до сих пор непонятен. Не только из-за времени требующегося на реагирование фоторецепторов, но и потому, что вычислительные процессы требуют

большого количества времени на обработку информации, а также на выбор правильного решения на основе результатов вычислений. Наивный процесс интеграции неосведомленный об изменениях о положении глаза потеряет информацию о мелких деталях путем усреднения.

8. Эксперименты

В качестве эксперимента, для лучшего и более наглядного понимания того, как статически видит глаз, была написана программа, демонстрирующая это. Программа писалась на языке JavaScript, что позволяет поместить её в интернет. На странице <https://kud8.github.io/> можно её увидеть. Для просмотра нужен современный браузер (Firefox 4+, Google Chrome 9+, Opera 12+, Safari 5.1+, Internet Explorer 11+). Программа состоит из двух подпрограмм:

1. Первая подпрограмма применяет все преобразования, которые будут описаны ниже, к исходному изображению, его можно скачать и передать дальше. Так как некоторые коллеги занимаются задачами, являющимися продолжением или наоборот предшественниками для нашей задачи, а именно – учёт движения глаз на картинке, поступающей на сетчатку и стабилизация видео (видео, в данном случае, это по сути всего лишь ряд изображений), создаваемое глазом. Поэтому и нужна возможность сохранить изображение.

2. Вторая подпрограмма накладывает изображение, полученное с помощью первой подпрограммы, на полусферу. В левом верхнем углу для сравнения можно увидеть исходное изображение. Получившуюся полусферу можно вращать с помощью мышки – достаточно лишь нажать на неё и удерживая мышку водить ею. В результате полусфера будет вращаться в трёхмерном пространстве вслед за мышкой. Также с помощью колёсика мышки можно приближать или отдалять эту полусферу.

Из-за того, что на периферии сетчатки в нашем глазу преобладают палочки, кото-

рые не могут воспринимать цвет, придётся заменять в этой области цветные пиксели изображения. Встаёт вопрос, на что заменять rgb каналы. Существует формула, дающая на этот вопрос ответ (Luma conversion formula) [15]. Она утверждает, что для определения яркости пикселя на картинке (как раз яркость палочки и воспринимают), следует использовать следующую формулу: $0.299 * r + 0.587 * g + 0.114 * b$.

В зависимости от того, как далеко находится пиксель от центра мы определяем, нужно ли его оставлять цветным или нет. В центральной окружности небольшого радиуса все пиксели цветные, однако из-за эффекта жёлтого пятна приходится немного изменять r и g каналы, так как жёлтый цвет – это как раз сочетание лишь красного и зелёного канала. После выхода из этой окружности красим пиксели с вероятностью, обратно пропорциональной расстояниям по абсциссе и ординате от центра изображения до рассматриваемой точки.

Чем ближе к краю сетчатки, тем меньшее количество фоторецепторов, в особенности палочек, имеется на ней. Таким образом количество палочек падает и ухудшается качество изображения – одна палочка принимает сигнал от большего по площади участка внешнего мира. Переноса это на терминины двумерного массива изображения получается, что нужно уменьшать качество изображения с помощью объединения пикселей.

"Темное пятно" представляет собой круг, радиусом $1/11$ от высоты картинки, так как он имеет диаметр примерно в 2 мм, в то время как вся сетчатка имеет радиус порядка 22мм. Определить его местонахождение тоже не сложно – под углом в 15 градусов от центра и примерно $1/8$ по абсциссе от края. Таким образом, если считать за начало координат левый верхний угол, то её ординату из определения тангенса можно определить по формуле: $y - tg(15) * (3/8) * w$, где y – ордината центра изображения и w – ширина (=высота) изображения.

Для наложения изображения, полученного первой подпрограммой, на трёхмерную полусферу использовалась библиотека `three.js`. Она как раз и используется для отображения анимированной компьютерной 3D графики. Благодаря использованию технологии WebGL мы можем выполнять часть кода на видеокартах, написав свои шейдеры на языке GLSL. Это значительно ускоряет отрисовку этой 3D графики.

9. Выводы

Человеческий мозг – удивительная вещь. Понимание, как он получает чёткое и объёмное изображение из того, что получает неидеальный глаз, до сих пор остаётся загадкой, которую ещё предстоит разгадать.

С увеличением производительности современных компьютеров и быстрым развитием машинного обучения, стремящегося работать как человеческий мозг, есть все предпосылки и условия для будущего развития математических моделей работы мозга, в том числе и обработке изображений. А с развитием и увеличением точности медицинских инструментов и приборов возможно улучшение нашего понимания того, как работает глаз и мозг.

Процесс движения глаза человека исследуется уже давно. Большой прорыв в исследованиях по этой теме произошёл в середине XX века. В работе коллеги, Игоря Титова [16], была проведена работа с алгоритмами пошагового построения движения глаза во время фиксации, однако там нет экспериментальное сравнение с реальным движением глаза.

Стабилизация изображения мозгом является очень сложной как с алгоритмической, так и с программистской точек зрения, однако неплохая работа проделана коллегой, А.С.Саликовым [17].

Огромное значение играет понимание того, как наш мозг преобразует поступающую ему информацию, когда глаз находится в

статическом положении. Решение этих проблем необходимо для построения математической работы нашего мозга при восприятии визуальной информации. Также нельзя забывать и о движении глаз и, конечно, о стабилизации всех этих неидеальных изображений мозгом.

10. Планы на будущее

В будущем предстоит объединить эти этапы, исправить ошибки и недочёты коллег. Также, перспективным является создание единой программы, которая объединяет и визуализирует все эти этапы. Причём главное, чтобы эта программа не только правильно работала, но и ещё быстро и оптимально строила изображения и видео.

Это нужно для того, чтобы была возможность работать с этим приложением в реальном времени. Это может помочь в перспективной и быстро развивающейся сфере компьютерного зрения и обработки изображений, ведь таким образом можно ближе подойти к пониманию того, как вместе работают наши глаза и мозг и как представить это в программном виде.

Сложной и одновременно нужной и перспективной задачей является стабилизация изображения, которая происходит в нашем мозгу. Дальнейшие исследования скорее всего и будут касаться именно этой задачи и её решения с точки зрения программирования.

Литература

- [1] H.R.Schiffman, Sensation and Perception – An Integrated Approach, fifth edition, 2003.
- [2] D.H.Hubel Eye, brain and vision, 1990.
- [3] Herbert Gross, Human eye, Handbook of Optical Systems: Vol. 4 Survey of Optical Instruments, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, 2008.
- [4] Austin Roorda, Human visual system-image formation, 539-555.
- [5] H. Helmholtz and J. P. C. Southall, eds., Helmholtz's Treatise on Physiological Optics, Optical Society of America, Rochester, 1924.
- [6] John G. Nicholls, A. Robert Martin, David A. Brown, Mathew E. Diamond, David A. Weisblat, Paul A. Fuchs. From Neuron to Brain 5th Edition

- [7] Charles A. Poynton, *Digital Video and HDTV: Algorithms and Interfaces*, Morgan–Kaufmann, 2003
- [8] D. Purves, G. J. Augustine, D. Fitzpatrick, W. C. Hall, A. LaMantia, J. McNamara, M. Williams, "NEUROSCIENCE: Third Edition"// Sinauer Associates, 2004. стр. 453 - 469
- [9] А.Л. Ярбус, "Роль движения глаз в процессе зрения"// Москва, 1965. стр. 7 - 12, 76-161
- [10] R. Engbert, K. Mergenthaler, P. Sinn, A. Pikovskiyb, "An integrated model of fixational eye movements and microsaccades"// PNAS USA, 2011.
- [11] R. Engbert, R. Kliegl, "Microsaccades Keep the Eyes' Balance During Fixation"// American Psychological Society, 2004.
- [12] Olveczky BP, Baccus SA, Meister M (2003) Segregation of object and background motion in the retina. *Nature* 423:401–408
- [13] Eric Whitmire, Laura Trutoiu, Robert Cavin, David Perek, Brian Scally, James O. Phillips, Shwetak Patel. *EyeContact: Scleral Coil Eye Tracking for Virtual Reality*, 2016.
- [14] Keith Rayner, Monica S. Castelhana. Eye movements during reading, scene perception, visual search, and while looking at print advertisements, 2006
- [15] Charles Poynton, "YUV and luminance considered harmful: a plea for precise terminology in video 2003
- [16] И.А. Титов, В.П. Алешин. Экспериментальное исследование цепочки "визуальное восприятие - действие" в визуально силовых системах. Международная конференция СРТ2017
- [17] А.С. Саликов, В.П. Алешин. Обратные задачи "ощущение-восприятие-действие" в виртуальном окружении. Международная конференция СРТ2017

Мультиинструментальный интерфейс для работы с цифровыми голограммами в виртуальном окружении (виртуальной реальности)*

В. А. Киселев, В. А. Пигарёв

kiselev.va@phystech.edu | vladislav.pigarev@phystech.edu

Московский физико-технический институт (государственный университет)

Работа нацелена показать актуальность виртуальной реальности, как средства взаимодействия с данными получаемыми методами цифровой голографической микроскопии. Ранее мы разработали средство для визуализации цифровых голограмм в виртуальной реальности. Сейчас сделав небольшой обзор методов записи цифровых голограмм, мы предприняли попытку спроектировать интерфейс для работы в виртуальном окружении с такими данными.

Ключевые слова: *Цифровая голография, точность, интерференция, строение организмов, виртуальная реальность, типизация, данные, массивы.*

Multi-tool interface for working with digital holograms in a virtual environment*

V. A. Kiselev, V. A. Pigaryov

Moscow Institute of Physics and Technology (State University)

The work is showed the relevance of virtual reality, as a means of working with a product accessible by digital holographic microscopy. Previously, we developed a tool for visualizing digital holograms in virtual reality. Now making a brief review of the methods of recording digital holograms, we made an attempt to create a system for working in a virtual environment with such data.

Keywords: *Three-dimensional visualization, precision, interference, structure of organism, virtual reality, typing, data, arrays.*

Введение

Голография — это область науки, изучающая технологии точной записи, воспроизведения и перестроения волновых полей оптического электромагнитного излучения. Голограммой называют зарегистрированное полное волновое поле объекта, т. е. запись как амплитудной, так и фазовой характеристики. В самом простом случае средой регистрации может выступать фотоплёнка. Под цифровой голографией стоит понимать способ регистрации волнового поля цифровым приёмником. Восстановление информации из цифровых голограмм в таком случае происходит с помощью численных методов.

Одной из перспективных областей применения цифровой голографии, а именно цифровой голографической микроскопии является наблюдение биологических объектов

и других микрочастиц. Визуализация цифровых голограмм биообъектов средствами виртуального окружения (виртуальной реальности) представляется нам вполне естественной: голограмма физически содержит информацию об объеме голографируемого объекта, а виртуальная реальность задумана быть объемной. В прошлом нами был разработан комплекс для визуализации цифровых голограмм биообъектов в среде виртуального окружения [1]. Тогда были реализованы следующие функции: перемещение точки наблюдения по трем пространственным координатам, выбор направления угла обзора, масштабирование размеров модели микрообъекта, кадрирование — определение размеров сцены, детализация — управление количеством полигонов, измерение расстояний между точками объекта, управление освещением в сцене.

Увеличение разрешения

В голографии разрешение оптических систем зачастую ограничивается числовой

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 17-57-52011 и 18-07-20030, 18-37-10008.

апертурой. Из-за её конечности удаётся зарегистрировать лишь низкочастотные волны. Чтобы это исправить, применяются разные методы по увеличению числовой апертуры приёмника. Их можно разделить на три основные группы.

В первой группе способы основываются на том, чтобы совершать несколько снимков ПЗС-камерой, пока та движется относительно исследуемого объекта. Траектория движения камеры такова, чтобы можно было зарегистрировать спектр под более широкими углами. В дальнейшем полученные изображения совместно обрабатываются и в результате получается одна цифровая голограмма высокого качества.

Примером подобной записи голограмм может служить опыт Дж. Х. Массига [2] и его команды. Ему удалось увеличить числовую апертуру, сделав и объединив 9 снимков с ПЗС-камеры, которую в процессе перемещал между разными позициями.

Также для примера подходит и опыт Л. Д. Мартинеза и его коллег [3], в котором они делали снимки, сдвигая камеру лишь на несколько микронов, что позволило не только увеличить разрешение, но и добиться высокой точности за счёт большой выборки. По отдельности голограммы могут иметь некоторую случайную зашумлённость, однако при объединении множества голограмм этот шум усредняется и получается более качественная картинка.

Во второй же группе в отличие от первой расположение элементов меняются для того, чтобы запечатлеть дифракционное поле освещённого объекта под разными углами. Регистрирующая система при этом остаётся в постоянном состоянии.

Например, команда С. А. Александрова [4, 5] в своём опыте вращала лишь регистрируемый предмет. Затем она сшила полученные снимки и получила новую искусственную апертуру, преодолев дифракционный предел системы.

Ю. О. Кузнецова с коллегами выполнила подобный опыт [6], однако она враща-

ла исследуемый объект относительно оптических осей. Так она перенаправила лучи, рассеянные под широким углом, в апертуру оптической системы и таким образом преодолела дифракционный предел.

Команда В. Мико использовала наклонное освещение [7, 8]. В одном опыте Мико использовал линейные лазеры с вертикальным резонатором [7]. Они позволили параллельно передавать несколько спектральных полос объекта путём использования пространственного мультиплексирования. В другом же опыте для улучшения разрешения Мико использовал спектральное смещение, вызванное наклонным освещением [8].

Использование в схеме дифракционной решётки

Во всех предыдущих опытах объединялись Фурье разложения спектров нескольких отдельных голограмм, для чего обычно приходится прибегнуть к сложным вычислениям. В третьей же группе методов для увеличения разрешения системы прибегают к использованию дифракционной решётки в оптической схеме, что позволяет избежать работы с Фурье разложением спектра.

Принцип действия метода довольно прост и эффективен. По сути дифракционная решётка позволяет перенаправить на ПЗС-камеру информацию, которая иначе бы пропала. В итоге получается, что на одной и той же матрице фиксируется 3 изображения. Можно получить сверхрешенные изображения при цифровой реконструкции этих мультиплексированных цифровых голограмм, путём увеличения численной апертуры в 3 раза.

Не любая дифракционная решётка может дать такое увеличение численной апертуры. Это становится возможным, когда она удовлетворяет трём важным свойствам.

Во-первых, она должна представлять собой двумерную шестиугольную структуру, что позволяет ей получать сверхразрешение в двух измерениях.

Во-вторых, дифракционная решётка должна быть фазовой, а не амплитудной. Амплитудные решётки не столь эффективны, так как для формирования изображения используется лишь та малая часть света, которая прошла через небольшие отверстия.

В-третьих, она должны быть изготовлена из электрооптической подложки. В таком случае путём изменения дифракционной эффективности можно будет настроить относительную интенсивность голограмм для улучшения качества конечной голограммы высокого разрешения.

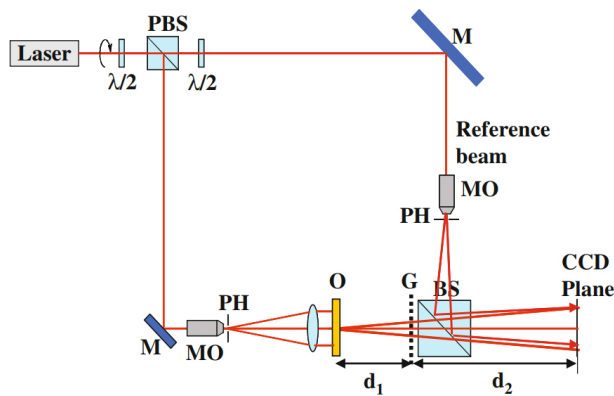


Рис. 1: Внеосевая схема Фурье с использованием дифракционной решётки для записи голограмм высокого разрешения

Концепция сцены виртуального окружения

Таким образом, голографические методы развиваются, а значит и мы должны развивать интерфейс виртуальной реальности для работы с ними, чтобы рано или поздно создать готовый к использованию исследователями инструмент. Его главной особенностью должна быть универсальность использования, ведь, как показывает практика, фотоприёмниками волнового фронта являются матрицы, а значит на вход программной среде поступает двумерный числовой массив.

Таким образом, были сформулированы требования, приведённые здесь в виде следующих формальных разделов — каждый раздел, описывает принципиальный режим в

рамках которого работает оператор в проектируемой нами среде ВО (ВР):

- Работа с сырыми исходными данными: Работа всегда начинается с пустого файла сцены, в котором есть набор инструментов. Один из инструментов позволяет загружать исходные данные цифровой голограммы. В функциональном меню должны быть выбраны параметры для построения модели (например в виде указания параметров записи голограммы). Восстановив голограмму оператор может делать с ней ряд операций по редактированию (требует дополнительного исследования), но в конечном счёте эта модель должна быть сохранена.
- Работа с восстановленными моделями: Подразумевает загрузку (или создание пустого) файла сцены, в которой содержится модель (или модели) с сохранёнными параметрами. Также в виртуальное пространство сцены могут быть добавлены новые модели объектов.
- Работа в режиме реального времени: Работа в режиме реального времени предполагает потоковое получение голограмм с цифровой фотоматрицы, и последующее численное восстановление и построение моделей. Требования временного разрешения вышеописанного цикла заданы параметрами интервала от 10 до 24 раз в секунду. Полезной была бы возможность записи сохранения кадра из режима реального времени в виде модели, а также запись её трансформации.
- Режим визуализации стереопары восстановленной с голограммы: Предполагает загрузку исходных данных цифровой голограммы и получение стереопары, путём численного восстановления для соответствующих углов наблюдения (освещение различными опорными лучами образующими стереопару).

Кроме того виртуальную сцену для работы с моделями необходимо наполнить интер-

фейсом инструментов со следующим функционалом:

- загрузка модели из файла/
- перемещение объекта;
- вращение объекта — позволяет вращать объект относительно заданной точки;
- свободная трансформация объекта — позволяет трансформировать объект, по осям x, y и z;
- линейка — позволяет измерять расстояния между двумя точками на одном объекте;
- управление освещением — позволяет позиционировать точечный источник освещения;
- перерисовка модели объекта с использованием другого количества полигонов;
- визуализация спектральных характеристик;
- инверсия оптического профиля относительно оси z.

Таким образом, требования к рабочей среде виртуального окружения сформировали ряд взаимосвязанных вопросов:

- Какие аппаратные и программные средства будут использованы?
- Как представлен пользовательский интерфейс?
- Как происходит хранение и обработка файловых данных?
- Каким методом происходит построение модели восстановленной цифровой голограммы?
- Позволит ли аппаратная составляющая реализовать требования временного разрешения для работы в режиме реального времени?

Аппаратные и программные средства

Очки виртуальной реальности, ПК и контроллер человеко-компьютерного взаимодействия для манипуляции со средой формируют аппаратную часть системы виртуального окружения. Для поставленных целей требуются вычислительные мощности, организующие быстрый канал передачи данных. В качестве очков подходят

oculus rift cv1 [9], а в качестве контроллера Leap Motion [10]. Последний позволяет осуществлять интерактивный трекинг мелкой моторики рук. Так же следует отталкиваться от технических характеристики ПК не ниже приведённых: процессор intel i7 4470 ГГц, видеокарта Geforce GTX 970, 16 Гб оперативной памяти, ssd. Так как аппаратная часть может быть заменена, то программный инструментарий должен быть легко портируем, чему соответствует мультиплатформенный движок Unity, с языком программирования C#.

Ниже мы приводим описание взаимодействия для каждого инструмента, которое по нашему мнению является оптимальным:

- Перемещение объекта:
Пользователь «захватывает» объект и перемещается с ним по сцене, выбирая тем самым положение для объекта. Окончание выбора по условному действию пользователя.
- Вращение:
Представлено в двух режимах: режим вращения и режим выбора оси вращения. Моделями кистей рук пользователь может вращать куб, вектор поворота которого присваивается объекту. Выбор оси происходит проецированием луча на объект по центральной линии поля зрения пользователя (метод ScreenPointToRay в unity). Окончание выбора и переключение между режимами по условному действию пользователя.
- Трансформация:
Происходит по трём локальным осям. Моделями кистей рук, можно регулировать соответствующий трек-бар. Окончание выбора по условному действию пользователя.
- Кадрирование:
Позволяет выбрать прямоугольную область, посредством размещения точек, соответствующих противоположным углам. Выбор происходит методом ScreenPointToRay, а окончание выбора

влечёт за собой удаление части объекта не попавшей в прямоугольную область.

- Перерисовка на другое количество полигонов:

Моделями кистей рук, можно регулировать соответствующий трек-бар, имеющий минимум равный 4 и максимум соответствующий минимуму размера карты высот. Окончание выбора по условному действию пользователя.

- Визуализация спектральных характеристик:

Моделями кистей рук, можно регулировать четыре трек-бара. Первая пара соответствует минимуму и максимуму отображаемых значений спектральных характеристик для фазы. Вторая пара соответствует минимуму и максимуму отображаемых значений для амплитуды. Так же имеется Checkbox, при включении которого визуализируется суммарная фаза и амплитуда.

- Измерение расстояния:

Позволяет поставить две точки на объекте методом ScreenPointToRay. По условному действию пользователя, выводит информацию о расстоянии в двух единицах измерения: локальной и глобальной. Окончание работы с инструментом по условному действию пользователя.

- Установка освещения:

Пользователь перемещается по сцене, выбирая положение для точечного источника освещения. Окончание выбора по условному действию пользователя.

- Инвертирование:

При выборе перестраивает объект, инвертированный по оси z.

По завершению применения всех инструментов на сцене остаётся преобразованный объект, а в памяти преобразованный и первоначальный, без сохранения промежуточных результатов.

Структура хранения файловых данных

Из сформулированных требований следует, что разрабатываемый проект должен пони-



Рис. 2: Пример иллюстрирующий рабочий процесс в проектируемом меню

мать несколько типов файлов, обеспечивающих хранение следующим паттернам:

- Цифровая голограмма:

В файле хранится набор двумерных матриц интерференционных картин (тип float), их размер и соответствующий им диапазон длин волн.

- Сцена:

В файле хранятся локальные ссылки на модели объектов, соответствующие им параметры и другие настройки сцены.

- Модель объекта:

В файле хранится набор моделей объектов в виде карт высот (двумерных матриц float, каждое значение в которых, соответствует высоте), их размер, соответствующий им диапазон длин волн и амплитудная маска (RGB).

Мы считаем, что наиболее подходящим форматом представления данных будет JSON.

Выводы

В настоящее время методы цифровой голографической микроскопии нельзя применять к любым объектам. Дело в том, что, во-первых, часть объектов являются диффузными (то есть не структурированы рассеивающими), во-вторых, цифровая голография макрообъектов довольно сложно реализуема. Тем не менее, при работе с биообъектами возникает потребность ими манипулировать. Сталкиваясь в таких случаях с микроуровнем, довольно часто используют оптические ловушки (или лазерные

пинцеты) [11, 12, 13], когда же речь идёт о макроуровне, то применяют различные инструменты, в основе которых лежат механические подвижки.

Таким образом, актуальным является создание универсальной автоматизированной системы виртуальной реальности как для работы как с микро, так и с макроуровнем. Стоит отметить, что получение модели всего объекта и манипулирование уже моделью, а не объектом по прежнему остаётся крайне актуальной технологией [1]. Внедрение же такой технологии в повседневную практику исследователей полезно, когда модель биообъекта можно получать в режиме реального времени, что требует как развития голографических методов, так и методов численной обработки и автоматизации процессов.

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 17-57-52011 МНТ_а и 16-07-00413 А.

Литература

- [1] Каленков С.Г., Каленков Г.С., Киселев В.А., Сысоев Н.А., Клименко С.В., Heiden W. «Визуализация цифровых голограмм биообъектов в среде виртуального окружения» // Труды Международной научной конференции «Ситуационные центры и ИАС4i для мониторинга и безопасности» SCVRT1516, 21-24 ноября 2015-2016 г. Протвино-Москва, 2016: Изд. ИФТИ. — С. 291-294.
- [2] J.H. Massig, Digital off-axis holography with a synthetic aperture. *Optics Letters*, 27, 2179–2181, (2002)
- [3] L. Martínez-León, B. Javidi, Synthetic aperture single-exposure on-axis digital holography. *Optics Express* 16, 161–169 (2008)
- [4] S.A. Alexandrov, T.R. Hillman, T. Gutzler, D.D. Sampson, Synthetic aperture Fourier holographic optical microscopy. *Phys. Rev. Lett.* 97, 168102 (2006)
- [5] T.R. Hillman, T. Gutzler, S.A. Alexandrov, D.D. Sampson, High-resolution, wide-field object reconstruction with synthetic aperture Fourier holographic optical microscopy. *Optics Express* 17, 7873–7892 (2009)
- [6] Y. Kuznetsova, A. Neumann, S.R. Brueck, Imaging interferometric microscopy—approaching the linear systems limits of optical resolution. *Optics Express* 15, 6651–6663 (2007)
- [7] V. Mico, Z. Zalevsky, P. Garcia-Martinez, J. Garcia, Single-step superresolution by interferometric imaging. *Optics Express* 12, 2589–2596 (2004)
- [8] V. Mico, Z. Zalevsky, C. Ferreira, J. Garcia, Superresolution digital holographic microscopy for three-dimensional samples. *Opt. Express* 16, 19260–19270 (2008)
- [9] Документация для разработчиков об Oculus Rift: <https://developer.oculus.com/>
- [10] Официальная документация камеры leap motion Web: <https://developer.leapmotion.com/>
- [11] Sang-Hyuk Lee, “Optimal integration of wide field illumination and holographic optical tweezers for multimodal microscopy with ultimate flexibility and versatility”, *Optics Express* 26, 8049-8058 (2018), doi:10.1364/OE.26.008049
- [12] Ivo T. Leite, Sergey Turtaev, Xin Jiang, Martin Šiler, Alfred Cuschieri, Philip St. J. Russell, Tomáš Čížmár, “Three-dimensional holographic optical manipulation through a high-numerical-aperture soft-glass multimode fibre”, *Nature Photonics* 12, 1749-4893 (2018), doi:10.1038/s41566-017-0053-8
- [13] Robert Kuschmierz, Dirk Schubert, Nektarios Koukourakis, Jürgen Czarske, “Holographic endoscope based on coherent fiber bundles and adaptive optics (Conference Presentation)”, *Proc.SPIE* 10502, (2018), doi:10.1117/12.2286934

Перспективы разработки программно-аппаратного комплекса для оперирования биологических макрообъектов в среде виртуальной реальности*

Х. П. Тирас^{1,2}, *В. А. Киселев*³, *С. В. Клименко*⁴
 tiras1950@yandex.ru | kiselev.va@phystech.edu

Пушино, ¹Институт теоретической и экспериментальной биофизики РАН

²Пушинский естественно-научный институт

³Московский физико-технический институт (Государственный университет)

⁴Нижегородский государственный архитектурно-строительный университет

Разработано механическое устройство управляемое с персонального компьютера для проведения разрезов на биологических макрообъектах. Проблема проведения операций на планариях являлась предпосылкой к разработке такого устройства. В данной работе поднимается вопрос актуальности использования подобных устройств, как для исследования планарий, так и в целом. А так же перспективы создания среды виртуального окружения (виртуальной реальности) для проведения хирургических операций на макрообъектах.

Ключевые слова: виртуальное окружение, виртуальная реальность, arduino, планарии, микроскопия, оперирование биологических макрообъектов.

Prospects for the development of a software and hardware complex for operating biological macroobjects in a virtual reality environment*

H. P. Tiras^{1,2}, *V. A. Kiselev*³, *S. V. Klimenko*⁴

Pushchino, ¹Institute of Theoretical and Experimental Biophysics of the Russian Academy of Sciences (ITEB RAS)

²Pushchino State Institute of Natural Sciences

³Moscow Institute of Physics and Technology (State University)

⁴Nizhny Novgorod State University of Architecture and Civil Engineering;

A mechanical device has been developed that is controlled from a personal computer for carrying out sections on biological macroobjects. The problem of conducting operations on planarians was a prerequisite for the development of such a device. In this paper, the question of the relevance of the use of such devices is raised, both for the study of planarians, and in general. And also the prospects for creating a virtual virtual reality for performing surgical operations on macroobjects.

Keywords: virtual environment, virtual reality, arduino, planaria, microscopy, operation of biological macroobjects.

Актуальность

Одна из центральных проблем биофармацевтики: решение противоречия между потребностью рынка в большем числе современных лекарственных средств и практикой использования массы позвоночных животных в качестве биологических моделей в доклинических исследованиях. Одновременно, нарастает разрыв между доклиническим и клиническими исследованиями, базирующийся на различиях между геномом человека и грызунов, являющихся основными объектами доклинических исследова-

ний. В этой связи все более актуальным является запрос на разработку альтернативных биологических моделей, базирующихся на других, в то числе, беспозвоночных животных, клеточные и молекулярные тесты. В последний годы внимание привлекла хорошо известная в биологии тест-система на основе регенерации плоских червей – планарий. для контроля действия физических и химических факторов на клеточную пролиферацию. Этот интерес, в том числе развивается по мере расшифровки генома планарий и сравнения специфических генов планарий и человека. В частности, оказалось, что консервативная часть генома стволовых клеток человека, ответствен-

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 17-57-52011, 18-07-20030.

ная за пролиферацию и дифференцировку, высоко гомологична (до 90%) соответствующим генам планарий [1, 2]

За последнее десятилетие была создана база данных генома планарий [3], и, в частности, определено около 30 геномных участков, сходных с генами важнейших нейродегенеративных заболеваний человека [4].

Параллельно развивается биофизические подходы к регистрации регенерации планарий, основанный на динамической морфометрии процесса регенерации системой неинвазивных методов, в том числе, прижизненной компьютерной морфометрией и контролем сверхслабой фотонной эмиссии, или хемилюминесценции [5, 6]

Регенерация пресноводных плоских червей – планарий – давно используется как тест-система для альтернативного биотестирования лекарственных препаратов – регуляторов клеточной пролиферации [1]. Она основывалась на неинвазивной регистрации регенерации методом прижизненной компьютерной морфометрии (ПКМ) [7].

Метод ПКМ включает ряд стандартов и стандартных операционных процедур, обеспечивающих надежность и валидность получаемых результатов. В их число входит стандарт на экспериментальных животных – бесполоую лабораторную расу планарий *Girardia tigrina*, а также стандартные процедуры создания экспериментальной модели регенерации, а также видеорегистрации процесса отрастания утраченной головной части тела, получения ее цифрового изображения и его анализа. Эта система нуждается в стандартизации процесса создания модели регенерации – точного рассечения тела планарии по линии между ушками интактной планарии. Проблема этой процедуры в том, что ее необходимо проводить на движущейся планарии в поле зрения бинокулярного микроскопа. Оператор должен провести точный разрез, в тот момент, когда он следит за животным в поле зрения микроскопа. Это сопряжено с экспериментальными ошибками, связанны-

ми с невозможностью точного контроля позиции глаз оператора и скальпеля, который находится в операционном поле. Рассогласование команды от глаза оператора к руке, производящей разрез, не позволяет создать стандартную операционную процедуру операции. Эта проблема относится не только к операциям на планариях, но также к многим другим экспериментальным операциям, которые желательно проводить на подвижных объектах без обездвиживания. Для этого был разработан специальный инструмент для подобных операций на т.н. макрообъектах, размер которых позволяет работать с применением бинокулярного микроскопа.

Постановка задачи и реализация

Формально систему с которой предстоит работать можно представить как микроскоп, направленный на подвижный столик с кюветой, в которой находится исследуемый объект. Микроскоп имеет цифровой фотоприёмник соединённый с ПК. С этого же ПК происходит управление режущим инструментом. То есть задачу можно разделить на два этапа: создание режущего инструмента и создание интерфейса для управления с ПК. Режущий инструмент должен иметь возможность быть закреплённым на микроскопе, однако, во-первых люфт лезвия не должен оказаться больше расстояния от плоскости подвижного столика микроскопа до его объектива, а во вторых из-за разных конструкторских особенностей микроскопов должна быть предусмотрена юстировка положения лезвия. Мы предложили вариант, представленный на Рис. 1. Здесь важно, что лезвие отделено от основного механизма механического воздействия, а в качестве управления рычагом используется лесочная система, так чтобы ось лезвия может быть смещена при необходимости. Такая конструкция позволяет регулировать как выступ лезвия, так и его люфт. Важной особенностью являются выделенные на рисунке юстировочные упоры. Как раз их поворот позволяет менять в ко-

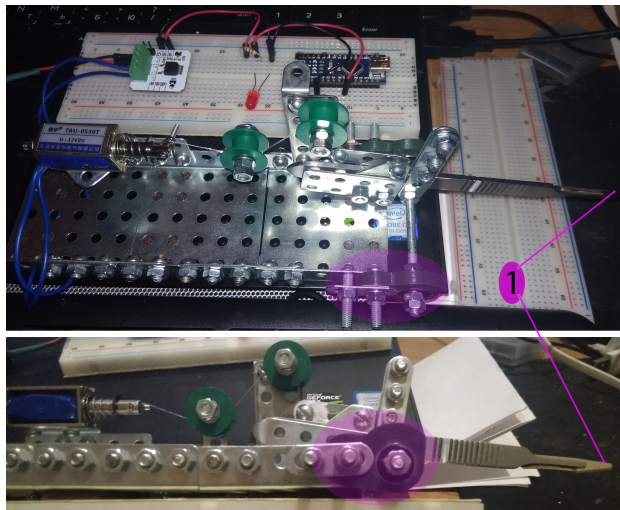


Рис. 1: Устройство для проведения хирургических операций на биологических макрообъектах управляемое с ПК. Юстировочный упор выделен подсветкой, 1 – место расположения лезвия

в конечном итоге угол наклона лезвия. Они были получены путём создания 3д модели и в последующем распечатаны на 3д принтере (Рис. 2). В качестве тянущей силы, приводящей к опусканию лезвия, был использован соленоид ТАУ-0530Т. Это довольно эффективное решение, однако силы тока выдаваемой контроллером в нашем случае оказалось мало, поэтому был задействован транзисторный аналог электромагнитного реле совместно с блоком питания на 2 [А] и 12 [В]. Управляется такая система контроллером arduino nano (v3) [8]. Чтобы не перегреть катушку соленоида на него отправляется разовый импульс, что заставляет лезвие опуститься, а затем упругая пружина возвращает систему в исходное состояние. Так же удалось значительно уменьшить задержку при взаимодействии с ПК – для этого была разработана программа, которая устанавливает связь с контроллером заранее, до начала операции. Протокол работы с контроллером нетривиален, и без подобного подхода значительную часть времени отнимал процесс налаживания связи между контроллером и ПК. То есть, наладив единожды канал связи с контроллером, тот переходит в режим ожидания пакета. При получении соответствующего пакета кон-



Рис. 2: Юстировочные упоры устройства для проведения хирургических операций: 3д модель и реализация на 3д принтере

троллер тут же отправляет импульс на соленоид.

Перспективы развития

Интерфейс взаимодействия с ПК на данный момент представлен набором стандартного софта идущим в комплекте к цифровой камере бинокулярного микроскопа, а так же реализованными на питоне скриптами — один заблаговременно устанавливает связь с контроллером, второй отправляет команду на контроллер в момент, когда нужно сделать разрез. Скрипт заранее настраивается на наиболее удобную клавишу, нажатие на которую и приводит к опусканию лезвия в конечном счёте. Однако, исследователь использующий микроскоп может либо сфокусироваться на объекте, либо на режущем инструменте. Это, порой, заставляет исследователя действовать вслепую (Рис. 3.). Разумеется, гра-



Рис. 3: Слева-сверху: чистое сфокусированное изображение с микроскопа. Справа-сверху: изображение с микроскопа сфокусированное на плоскости макрообъекта, со скальпелем над ним. Далее изображения покадрово демонстрируют перефокусировку на скальпель

фическая разметка подсказывающая в каком месте располагается лезвие – была бы тривиальным решением. Тем не менее, мы считаем что создание среды виртуального окружения (виртуальной реальности) намного актуальнее для работ подобного рода. Ведь исследователь, находящийся в среде виртуальной реальности может сосредоточиться на исследуемом явлении, и не отвлекаться на внешние факторы. Кроме того, ранее мы разработали инструмент для визуализации цифровых голограмм в виртуальной реальности [9], то есть позволили оператору наблюдать пространственную структуру биообъекта. Таким образом, примитивный интерфейс виртуальной реальности должен позволять оператору прицелиться (выбрать место для разреза). Таким образом, как минимум в виртуальную реальность нужно вывести изображение с цифровой камеры микроскопа, разместить там модель лезвия, и совместить VR среду с устройством ввода для возможности

отдавать команду для совершения разрезов. Отсканировав лезвие с большим разрешением и обведя в программе 3ds max контур можно задать толщину замкнутому сплайну и получить полигональную модель, что мы и сделали Рис. 4. Так как

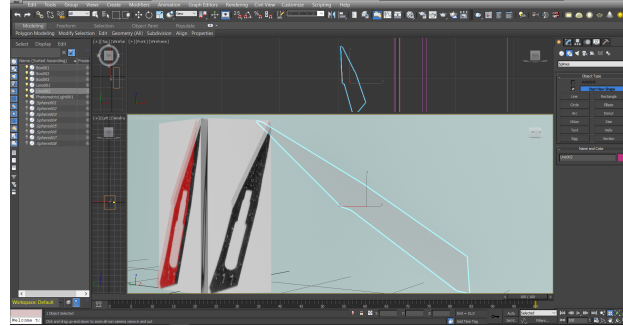


Рис. 4: Процесс получения полигональной модели в программе 3ds max

аппаратная часть будет дорабатываться, то программный инструментарий должен быть легко портируем, чему соответствует мультиплатформенный движок Unity 3d, поддерживающий язык программирования C#. Тогда для вывода изображения с камеры микроскопа можно использовать класс WebCamTexture.cs [10]. В качестве очков виртуальной реальности подходят oculus rift cv1 [11], а в качестве контроллера для взаимодействия со средой Leap Motion [12]. Последний позволяет осуществлять интерактивный трекинг мелкой моторики рук. Так же следует отталкиваться от технических характеристик ПК не ниже приведённых: процессор intel i7 4470 ГГц, видеокарта Geforce GTX 970, 16 Гб оперативной памяти, ssd.

Довольно перспективным направлением является трекинг макрообъектов для предсказания траектории их движения при проведении хирургических операций. Это позволило бы, как предсказывать траекторию движения биообъекта для повышения точности операции, так и производить сбор информации о биообъектах для последующего анализа. В этом смысле стоит рассматривать виртуальную реальность как возможность создания полноценной рабочей среды для исследователя, выходящей

далеко за рамки проведения хирургических разрезов. Кроме того, установка дополнительной камеры, отслеживающей перемещение биообъекта, совместно с разработкой алгоритмов по машинному зрению открывает возможности по увеличению точности подобного трекинга — ведь в связке с получением трёхмерной модели исследуемого биообъекта это позволит получать три координаты изменения траектории, вместо двух. Не менее актуальным являются возможности удалённого управления и группового взаимодействия. Когда подобная среда виртуального окружения превратится в виртуальную лабораторию, что позволит взаимодействовать группе исследователей из разных частей мира, проводящим совместный эксперимент над биообъектом.

Литература

- [1] Sánchez Alvarado A, Newmark PA, Robb SM, Juste R. The Schmidtea mediterranea database as a molecular resource for studying plathyhelminthes, stem cells and regeneration. // *Development*. 2002. V.129. N.24. P. 5659-5665.
- [2] Labbé RM, Irimia M, Currie KW, Lin A, Zhu SJ, Brown DD, Ross EJ, Voisin V, Bader GD, Blencowe BJ, Pearson BJ. A comparative transcriptomic analysis reveals conserved features of stem cell pluripotency in planarians and mammals. // *Stem Cells*. 2012. Aug; V.30. N.8. P.1734-1745. doi: 10.1002/stem.1144.
- [3] База данных генома планарий Web: <http://genome.wustl.edu/genomes>
- [4] Elliott, Alvarado, Elliott SA, Sánchez Alvarado A. The history and enduring contributions of planarians to the study of animal regeneration. // *Wiley Interdiscip Rev Dev Biol*. 2013 May-Jun; V.2. N.3. P.301-26. doi: 10.1002/wdev.82. Epub 2012 Jul 23
- [5] Тирас Х.П., Асланиди К.Б. Регламентация условий культивирования планарий и параметров морфометрического эксперимента // *Современные проблемы науки и образования*. 2016. № 6.
- [6] Х.П.Тирас, С.В.Гудков, В.И.Емельяненко, К.Б.Асланиди. Собственная хемиллюминесценция необластов планарии в процессе регенерации. // *Биофизика*. 2015. Т.60. №5. С. 975-980
- [7] Тирас Х.П., Асланиди К.Б. Тест-система для доклинического исследования медицинской и экологической безопасности на основе регенерации планарий // *ТулГУ*, 2013. С. 63., Тирас Х.П., Асланиди К.Б. 2016
- [8] Официальная документация Arduino Web: <https://www.arduino.cc/en/Guide/HomePage>
- [9] Каленков С.Г., Каленков Г.С., Киселев В.А., Сысоев Н.А., Клименко С.В., Heiden W. «Визуализация цифровых голограмм биообъектов в среде виртуального окружения» // *Труды Международной научной конференции «Ситуационные центры и ИАС4i для мониторинга и безопасности» SCVRT1516*, 21-24 ноября 2015-2016 г. Протвино-Москва, 2016: Изд. ИФТИ. — С. 291-294.
- [10] Официальная документация Unity 3d. Класс WebCamTexture: Web: <https://docs.unity3d.com/ScriptReference/WebCamTexture.html>
- [11] Документация для разработчиков об Oculus Rift: <https://developer.oculus.com/>
- [12] Официальная документация камеры leap motion Web: <https://developer.leapmotion.com/>

Современные методы построения трехмерной модели операционного поля по видеоряду*

С. В. Клименко¹, **В. И. Пестриков²**, **А. М. Чернецкий²**
 pestrikov@phystech.edu | arkadiy.chernetskiy@phystech.edu

¹Институт физико-технической информатики, Протвино, Россия

²Московский физико-технический институт (государственный университет),
 кафедра физико-технической информатики

Цель данной работы — обзорный анализ современных методов трехмерной реконструкции операционного поля в рамках малоинвазивной хирургии. В работе представлен обзор некоторых современных решений из семейства алгоритмов SLAM, а также представлены работы исследователей по изучению эффективности использования нейронных сетей в рамках поставленной задачи. Немаловажным пунктом является исследование существующих наборов экспериментальных данных, необходимых для обучения нейронных сетей и тестирования алгоритмов в целом.

Ключевые слова: Минимально инвазивная хирургия, structure from motion, SfM, simultaneous localization and mapping, SLAM, 3D реконструкция, нейронные сети, глубокие нейронные сети.

Modern methods for constructing a three-dimensional model of the operating field from the video sequence*

S. V. Klimenko¹, **V. I. Pestrikov²**, **A. M. Chernetskiy²**

¹Institute of Computing for Physics and Technology

²Moscow Institute of Physics and Technology (State University),
 Department of Physics and Technology

The subject of this work is the review of modern methods of 3-dimensional reconstruction of the operating field in the framework of minimally invasive surgery. The work presents the overview of certain modern solutions from the family of SLAM algorithms, as well as works of researchers focused on the study of the efficiency of neural networks applied to the given task. An examination of existing sets of experimental data necessary for training neural networks and testing algorithms in general is also an important point of the current paper.

Key words: Minimal invasive surgery, structure from motion, SFM, simultaneous localization and mapping, SLAM, 3D reconstruction, neural networks, deep neural networks.

Введение

Компьютерные технологии стали практически неотделимы от жизни современного человека. В частности, использование технологий визуализации данных позволяет получать информацию в более наглядном и интуитивно понятном виде, что значительно упрощает процесс принятия решений и повышает эффективность работы. Технологии обработки изображений помогают создавать алгоритмы, способные самостоятельно принимать решения на основе полученных данных, либо помогать пользователю в различных ситуациях, таких как управление автомобилем или фотографирование пейзажа.

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 16-37-50054, 18-07-20030, 18-37-10008.

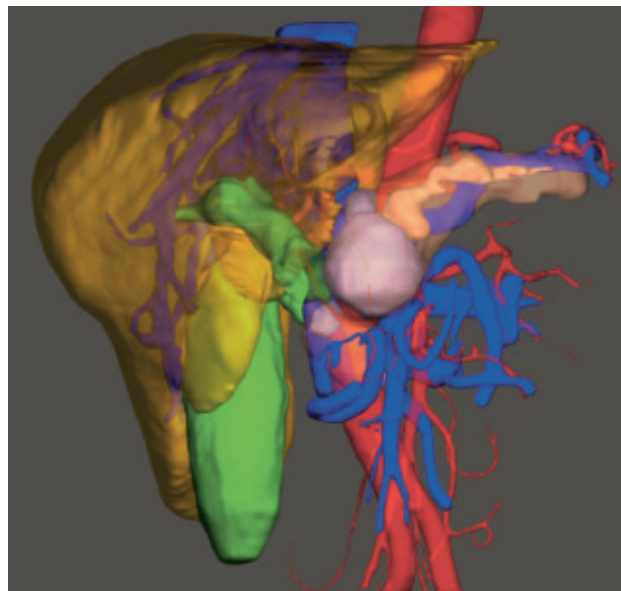


Рис. 1: Полигональная 3D-модель до операции

Хирургия также не осталась в стороне, и на сегодняшний день множество исследо-

вательских команд стараются помочь хирургам увеличить эффективность их работы путем внедрения решений, использующих подобные технологии визуализации. Многие данные медицинских исследований, например результаты томографии, могут быть представлены в более наглядном и удобном для анализа виде. Например, исследовательская группа специалистов с Самарского государственного медицинского университета [1] на основе результатов томографии смогла создать подробную трехмерную модель операционного поля, которую возможно использовать на этапе планирования операции (Рис. 1). По статистике, предоставленной в статье, эффективность хирургического персонала при использовании данного инструмента увеличилась: время операции уменьшилось на 11,5%; благодаря подготовке на основе созданной модели, во время операции удалось избежать перекрытия основного кровотока у большинства пациентов; показатели операционной кровопотери уменьшились. В свою очередь, сами хирурги отметили удобство предлагаемой системы, повышение уверенности в исходе операций за счет упрощения подготовки и увеличения предсказуемости осложнений. Впоследствии, на основе данного инструмента был разработан комплекс «Автоплан» — система, способная построить трехмерную модель операционного поля на основе данных томографии и выводить ее в системе дополненной реальности хирургам непосредственно во время проведения малоинвазивных хирургических операций. Дополнительная информация позволяет лучше спланировать предстоящие действия, а также лучше ориентироваться во время проведения хирургического вмешательства (Рис. 2).

Ещё одним вариантом применения системы трехмерного моделирования операционного поля является создание тренажёра с использованием виртуальной реальности в целях обучения будущих хирургов. С учётом достаточно стремительного раз-

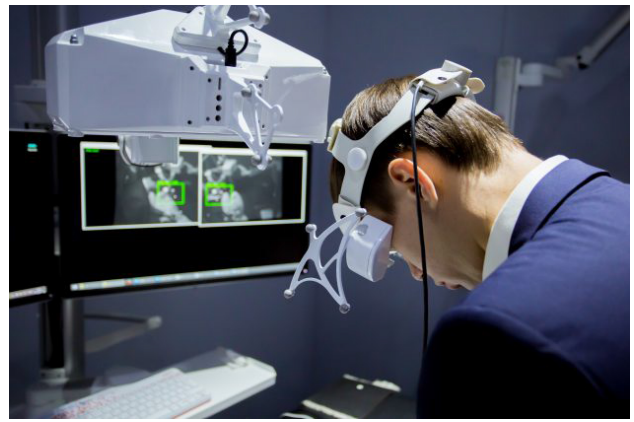


Рис. 2: Демонстрация системы «Автоплан»

вития новых технологий, необходимо создавать и использовать новые инструменты и методы обучения будущих специалистов. В качестве примера работы по данным направлениям можно привести такие системы, как тренажёры *ArthroS* [2] или отвлечение от разработки системы «Автоплан» — «*Surgera VR*». Данные системы уже сейчас показывают неплохие результаты, хотя и используют в большей степени синтетические изображения, построенные по данным МРТ.

Последний пункт необходимо осветить детальнее. Несмотря на все преимущества использования МРТ, данный метод имеет существенные недостатки. Во-первых, на сегодняшний день далеко не в каждой больнице и не для каждого пациента удаётся сделать томографию. Во-вторых, изображения, а, следовательно, и модели, полученные с помощью МРТ, не имеют текстуры и не могут передать реальные цвета, что не позволяет увидеть полную картину того, что реально происходит в операционном поле. Это может существенно отразиться как на качестве планирования операции, так и на качестве симуляции тренажёров. Для решения данной проблемы можно использовать системы реконструкции трехмерных моделей из видеоряда, в качестве которого может использоваться видео, полученное с эндоскопа или камеры хирургического робота.

Для этого случая, в рамках малоинвазивной хирургии ведутся исследования по

созданию алгоритма, способного с помощью видеоряда с камеры, без использования дополнительных датчиков, воссоздать трехмерную модель операционного поля и, более того, ориентироваться в построенной модели с минимальной ошибкой. Одними из первых алгоритмов, решающих данную задачу, были алгоритмы *Structure From Motion (SfM)*. Впоследствии к ним на замену пришли алгоритмы *Simultaneous Localization and Mapping (SLAM)*. Наиболее популярными оказались алгоритмы, основанные на поиске особых точек. В частности, можно отметить алгоритм ORB-SLAM, показывающий лучшие результаты среди алгоритмов своего семейства. В то же время, работ по использованию другого типа алгоритмов SLAM - Dense SLAM довольно мало, хотя всё же можно отметить работу [3] из института интеллектуальных систем Макса Планка, в которой описаны преимущества алгоритмов семейства Dense SLAM в свете выполнения поставленной задачи.

Одной из важных особенностей при работе с видеорядом с хирургических операций является подвижность объектов, находящихся в кадре. Живые ткани имеют хорошую эластичность, и, кроме того, могут двигаться и изменять свою форму, из-за чего алгоритмы, ориентированные на статичные модели, дают сбой. Во избежание этого, требуется особый подход, способный либо отделять подвижные части операционного поля от неподвижных для ориентации только по последним, либо позволять некоторую изменчивость модели, как, например, описано в статье [4].

За последний год набирают все большую популярность алгоритмы SLAM на основе нейронных сетей. Так, всё та же группа исследователей из института Макса Планка опубликовала ряд работ [5, 6], связанных с использованием рекуррентных и свёрточных нейронных сетей и выполняющих задачу по навигации миниатюрного роботизированного зонда внутри человеческого тела. Более подробно эти работы будут рас-

смотрены далее, пока же стоит отметить, что результаты экспериментов на тестовых данных выглядят оптимистично.

Отдельно стоит отметить проблему построения наборов данных для тестирования алгоритмов, их сравнения, а, в случае с нейронными сетями, и обучения алгоритмов. На сегодняшний день, исследователям, работающим с медицинскими данными, приходится создавать свои собственные наборы данных. К сожалению, полученные модели не всегда публикуются в свободном доступе.

Далее в статье подробнее рассматриваются вышеописанные подходы к построению трехмерных моделей операционного поля. В следующей главе описаны алгоритмы SfM и SLAM. Отдельно рассматриваются SLAM алгоритмы, использующие нейронные сети глубинного обучения. Приводится информация по существующим наборам данных для обучения и тестирования алгоритмов, работающих в сфере медицины. В конце работы приводится заключение и основные выводы.

Алгоритмы Structure from Motion и Simultaneous Localization and Mapping

Structure from Motion. Исторически первыми алгоритмами реконструкции трехмерных моделей из изображений являются Structure from Motion. Алгоритмы данного семейства используют серию изображений одного объекта, сделанных с разных ракурсов и, возможно, с различных камер. Сначала они анализируют входные изображения и выделяют на них характеристические точки. Далее, с помощью триангуляции находятся координаты этих точек в трехмерном пространстве. При этом возникает задача нахождения координат камер при съемке. При достаточном количестве найденных характеристических точек на изображениях можно составить необходимое количество уравнений для определения координат камер и построения модели объекта. Более подробно про работу алгоритмов SfM можно узнать из книги [7].

Одним из примеров реализации данного алгоритма может послужить работа исследователей из Центра по обработке медицинских изображений Лондона [8]. Данная работа использует известный алгоритм *Lucas-Kanade* для минимизации ошибки определения движения, после чего особые точки вычисляются как точки с минимальной разницей интенсивностей между изображениями. Главной особенностью данной работы является то, что, в случае малого количества или малого времени «активности» особых точек, авторы предпринимают попытки восстановить недостаток данных. Для этого используются методы, представленные в статьях [9, 10]. В дополнение к этому представлен алгоритм, улучшающий выбор данных для восполнения недостатка характеристических точек. Полученный алгоритм прошел тестирование на синтетических данных и видеозаписях работы робота «*da Vinci*». На результирующих данных (Рис. 3) можно увидеть увеличение качества работы представленного алгоритма относительно работы [9]. К сожалению, других алгоритмов для сравнения в статье представлено не было.

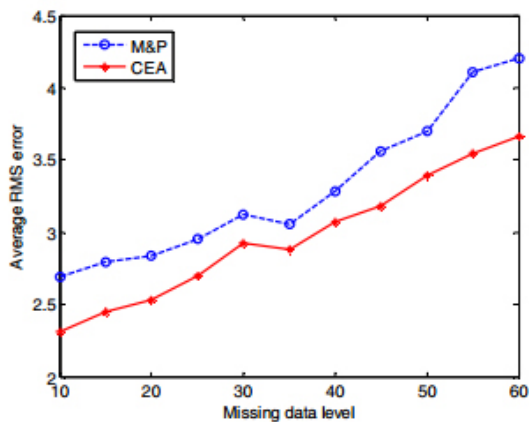


Рис. 3: Экспериментальные данные на синтетических тестах

Особенность семейства алгоритмов SfM заключается в том, что для корректной работы они требуют обработки сразу всей коллекции снимков, в то время как построение модели операционного поля в реальном времени требует возможности постоянного

обновления входящих данных. Из-за необходимости обновлять входящие данные (в частном случае — обрабатывать входящее видео покaдрово), были разработаны алгоритмы семейства **SLAM** — *Simultaneous Localization and Mapping* (одновременная локализация и построение карты).

Simultaneous Localization and Mapping.

Основной задачей алгоритмов SLAM является определение местоположения сенсора при перемещении в неизвестной среде по неизвестной траектории, а также построение карты среды, в которой происходит движение. В отличие от задачи SfM, для решения задачи SLAM может использоваться не только видеокамера, но и другие датчики. В контексте статьи необходимо сфокусироваться на подмножестве алгоритмов, называемым visual-SLAM или vSLAM, в рамках которых главными и единственными сенсорами являются видеокамеры (монокамеры или множественные объективы, объединённые в массив сенсоров). Главной отличительной особенностью vSLAM от SfM является то, что алгоритмы SLAM работают именно с видеопотоком, а не набором изображений, что существенно увеличивает количество информации, предоставляемое на вход алгоритму.

Важно отметить, что алгоритмы vSLAM (далее просто SLAM) разделяются на 4 категории по работе с данными: плотные и разреженные; прямые и косвенные [11] (Рис. 4). Особого внимания заслуживают прямые-плотные и косвенно-разреженные алгоритмы.

Косвенно-разреженные алгоритмы (или алгоритмы, основанные на поиске и обработке особых точек) похожи на алгоритмы SfM по своей структуре. Они используют информацию лишь о небольшом количестве зон на изображении вокруг особых точек для отслеживания движения сенсора. Наиболее ярким представителем данного вида алгоритмов является **ORB-SLAM** [12]. Данный алгоритм состоит из 3 параллельно работающих задач: трекинг положения сен-

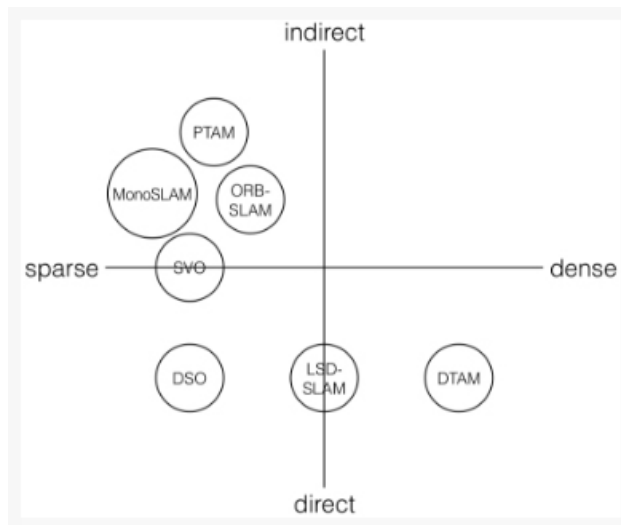


Рис. 4: Некоторые популярные алгоритмы SLAM

сора по особым точкам, предоставляемым алгоритмом ORB; построение модели среды, в которой перемещается сенсор; модуль, занимающийся поиском ключевых кадров для замыкания траектории. Последний модуль повышает эффективность работы алгоритма в условиях закрытого пространства, однако может также и ухудшить качество, если не учесть возможные деформации окружающего пространства.

Прямые-плотные (или просто прямые) алгоритмы SLAM стараются использовать информацию, полученную со всего изображения. В основном алгоритмы данной категории стараются строить и, впоследствии, модифицировать карты глубин по мере получения новых кадров. Среди примеров прямых алгоритмов можно отметить алгоритм **LSD-SLAM** [13]. LSD-SLAM состоит из 3 модулей: модуль трекинга, определяющий положение сенсора; модуль оценки карты глубины, который уточняет текущую карту глубины или заменяющий текущий ключевой кадр на новый; модуль, обновляющий карту окружающего пространства. Прямые алгоритмы имеют преимущество в работе с плохо текстурированными объектами, однако они более ресурсоёмкие в вычислениях и обладают меньшей робастностью относительно освещения.

ORB-SLAM. В рамках нашей задачи, несмотря на преимущество прямых алгоритмов, наибольшую известность получили именно алгоритмы, использующие особые точки. В частности, многие исследователи старались использовать ORB SLAM с некоторыми поправками. Так, в статье команды исследователей из института Бёрменса [14] представлен оптимизированный по параметрам алгоритм ORB-SLAM, а так же представлена методика для построения качественной трехмерной модели. Для тестирования полученных алгоритмов использовались синтетические данные с GroundTruth и данные Центра Лапароскопии Хамлина [15] (Рис. 5). В целом работа показывает качество применения стандартных алгоритмов SLAM с использованием особых точек в условиях малоинвазивной хирургии. В гораздо большей степени данная работа рассказывает, как получить качественную трехмерную модель операционного поля из скудного набора особых точек. Результаты работы показывают хороший уровень качества и достаточно оптимистичны. Однако, всё же следует помнить, что основное тестирование качества проводилось на синтетических данных, которые не всегда и не во всём соответствуют реальным.

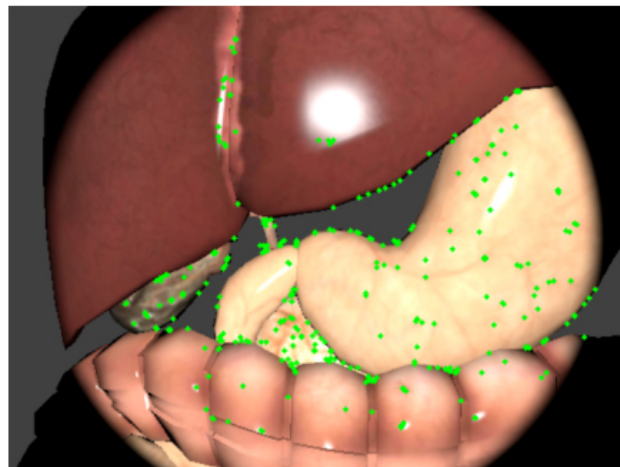


Рис. 5: Пример работы алгоритма ORB

В качестве другого примера использования ORB SLAM можно представить работу Технологического Университета Сиднея [4].

В ней исследователи предоставляют улучшенный алгоритм ORB SLAM, оптимизированный для решения проблемы быстрых движений, когда позиция камеры меняется слишком резко, что вызывает ошибки при определении местоположения камеры, а также создаёт размытые кадры, вносящие искажения в общую карту пространства. Также была представлена оптимизация работы с особыми точками. В статье предлагается производить отсев непригодных особых точек для операции замыкания, что позволяет уменьшить ошибку позиционирования, связанную с деформацией тканей. Кроме этого, в работе предлагается подход, использующий связку CPU+GPU для обработки данных, полученных от алгоритма ORB SLAM в реальном времени с целью построения трехмерной модели операционного поля. Главным преимуществом данного подхода является возможность обрабатывать движущиеся объекты в поле зрения и в реальном времени перестраивать трехмерную модель, что может позволить хирургам видеть актуальную высокополигональную модель операционного поля практически без задержек. Тестирование данного алгоритма проводилось на синтетических данных с GroundTruth и данных Центра Лапароскопии Хамлина [15] по аналогии с предыдущим. Полученные результаты показывают приемлемое качество. Однако, как отмечают авторы алгоритма, в работе всё так же имеются проблемы с локализацией при быстром движении, а также существуют проблемы с засветкой поверхности, особенно при сильном приближении к тканям, из-за чего в итоговой модели часто бывают проблемы с цветами и яркостью различных элементов.

LSD-SLAM. Прямые алгоритмы SLAM в редких случаях использовались для решения задач реконструкции в условиях малоинвазивной хирургии. Вероятно, основными причинами такого положения дел можно считать недостаточное итоговое качество работы или относительно малый возраст этого семейства алгоритмов. Приме-

ру удачного применения прямого алгоритма посвящена работа института интеллектуальных систем Макса Планка [3]. В своей статье авторы представляют алгоритм для трекинга и реконструкции операционного поля для капсульного роботизированного зонда, имеющего размер с обычную пилюлю. Несмотря на то, что источником видеопотока является не эндоскоп, а небольшая камера на борту зонда, задача практически идентична исследуемой с небольшими поправками на увеличение размера операционного поля. Можно отметить вероятное отсутствие резких движений робота и движение только в одном направлении, что делает практически бесполезным систему замыкания алгоритмов, работающих на основе особых точек. Сам алгоритм разбит на несколько небольших систем и является расширенным вариантом классического прямого алгоритма (Рис. 6). В основе лежит гибридная система из использования *Shape from Shading (Sfs)* [16] и прямого метода SLAM, что позволяет улучшить качество построения карты глубины для каждого кадра. Кроме того, каждый кадр проходит предобработку, очищающую его от бликов, отражений, а также уточняющую края. Ещё одним важным моментом является использование техники активных и неактивных зон. По своей сути данная система является аналогом системы замыкания, то есть активные зоны становятся неактивными, если не появляются некоторое время. Однако, если находится пересечение активных и неактивных зон, — происходит замыкание. К слову, также можно отметить, что алгоритм реализован на CUDA и работает с использованием GPU.

Для тестирования алгоритма авторы работы подготовили собственный датасет. Для этого зонд с камерой был пропущен через свиной кишечник. При этом местоположение и ориентация зонда отслеживались с помощью магнитных датчиков и оптической системы позиционирования «Optitrack». Как показывает таблица результатов (Рис. 7), точность итогового ал-

горитма достаточно высока. В среднем ошибка трекинга составляла около 2-3 см от Ground Truth. К сожалению, факт того, что был использован собственный набор данных, не даёт возможности оценить качество работы алгоритма относительно других.

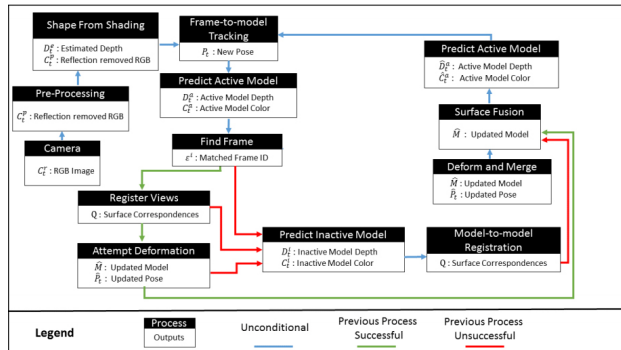


Рис. 6: Схема работы прямого алгоритма SLAM с использованием SfS

Trajectory ID	POTENSIC	VOYAGER	AWAIBA	LENGTH
1	0.015	0.018	0.020	0.414
2	0.018	0.021	0.023	0.513
3	0.017	0.019	0.025	0.432
4	0.032	0.035	0.042	0.478
5	0.035	0.038	0.045	0.462
6	0.038	0.041	0.048	0.481
7	0.041	0.043	0.049	0.468

Рис. 7: Среднеквадратичная ошибка позиционирования капсульного робота в метрах относительно Ground Truth

Алгоритмы SLAM на основе нейронных сетей

В последнее время доступность больших вычислительных ресурсов возродило интерес к искусственным нейронным сетям. На сегодняшний день вычисления на графических процессорах позволяют создавать сложные архитектуры нейронных сетей и конструировать так называемые глубинные нейронные сети. Исследователи в области компьютерного зрения стремятся переложить задачу по отбору признаков, необходимых для решения их задач, на нейронные сети, с чем те довольно неплохо справ-

ляются. Подробнее на принципах работы различных нейронных сетей мы останавливаться не будем, об этом можно прочитать во множестве общедоступных статей и источников [17, 18, 19]. Стоит отметить, что популярность нейронных сетей не обошла стороной и SLAM алгоритмы в малоинвазивной хирургии. Различные команды исследователей стараются заменять некоторые части алгоритмов глубинными нейронными сетями с целью улучшения качества и скорости работы.

В качестве одного из примеров можно привести команду исследователей из центра роботизированной хирургии Хамлина [20]. В их работе представлена самообучающаяся нейронная сеть, строящая карты глубины по видеоряду с хирургических операций, что является одной из ключевых задач SLAM алгоритмов. В своей работе авторы предложили сверточную сеть архитектуры кодер-декодер. В данной архитектуре начальные слои обрабатывают входящее изображение для выделения признаков, которые поступают во вторую половину сети для решения конкретной задачи. В базовой версии предложенной сети карта глубины строится для левого кадра. На основании полученной карты глубины из правого изображения реконструируется исходный кадр. Для оценки качества исходное изображение сравнивается с его реконструкцией. В статье авторы предложили улучшение алгоритма. Было предложено строить карту глубины для обоих кадров из пары. При этом в качестве меры ошибки использовалась сумма разниц между реконструируемыми изображениями и их оригиналом, а также несоответствие получившихся карт глубины. Тестирование проводилось на открытых датасетах этого же института [15]. По результатам работы алгоритма можно увидеть, что итоговая сеть достаточно качественно строит карты глубин по представленным изображениям. Более того, модификация алгоритма позволила значительно улучшить результат (Рис. 8).

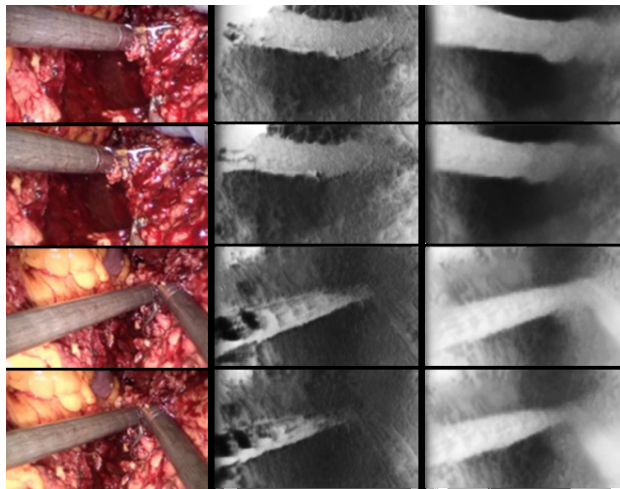


Рис. 8: Результат работы алгоритма [20]. Исходное изображение слева, Результат работы базового алгоритма по центру и результат работы модифицированного алгоритма справа.

Другая, уже упомянутая ранее, группа исследователей из института интеллектуальных систем Макса Планка представила работу [5], в которой решается не менее важная задача SLAM алгоритмов, а именно построение траектории движения датчика. Для решения данной проблемы представлена рекуррентная сверточная сеть. В качестве подготовки для каждого кадра из видеоряда с помощью алгоритма SfS строится карта глубины и объединяется с исходным кадром в одно RGBd изображение. На вход предложенной нейронной сети подается пара последовательных RGBd изображений из видеоряда. Первой частью сети выступают несколько сверточных слоев, которые выделяют ключевые признаки из поступивших кадров. Выделенные признаки подаются на рекуррентную часть нейронной сети, которая состоит из двух LSTM [21] слоев. На выходе нейронной сети получается вектор из 6 координат — три координаты смещения и три координаты поворота. В качестве функции ошибки использовалась L_2 норма. Исследователи протестировали предложенную сеть на различных данных и представили результаты, из которых видно, что предложенное решение значительно превосходит по точности классические алгоритмы, такие как ORB-SLAM и LSD-SLAM (Рис. 9).

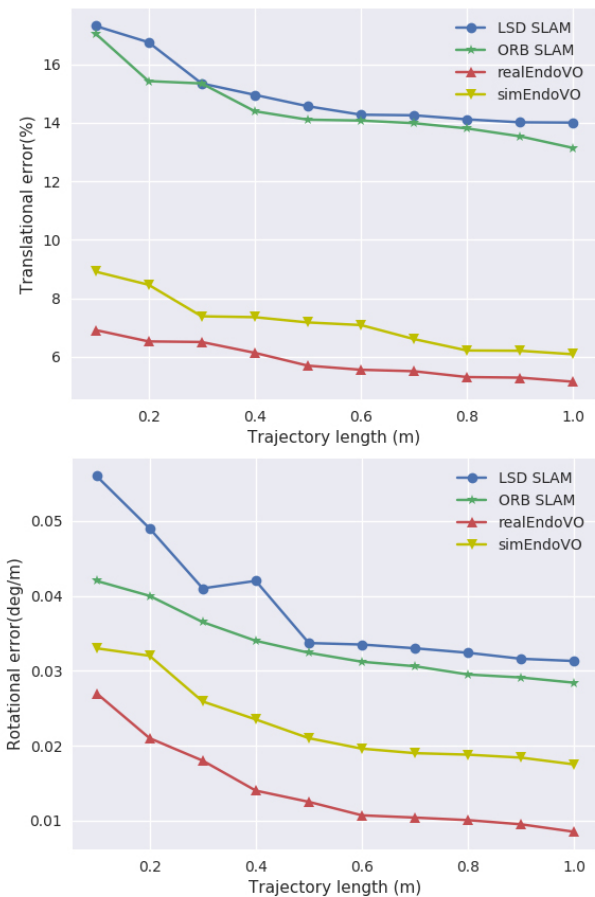


Рис. 9: Сравнение ошибок в определении сдвига (сверху) и поворота (снизу) алгоритмов EndoVO и классических SLAM алгоритмов

Представленные выше работы используют нейронные сети в различных частях алгоритмов. Логично было бы предположить, что их можно объединить, используя различные нейронные сети в различных частях алгоритма. Примером такого алгоритма может служить следующая работа [6]. В статье авторы представили две сверточные сети архитектуры кодер-декодер. При этом при обучении обе сети «помогают» друг другу и, таким образом, могут обучаться без учителя. Сеть для построения карты глубины принимает на вход одно изображение и выдает на выходе его карту глубины. Вторая сеть для вычисления положения камеры принимает на вход исходное изображение, а также предыдущий и следующий кадры и выдает шестимерный вектор смещения. На основе полученной карты глубины и данных о смещении камеры

реконструируются соседние кадры видеоряда. В качестве функции ошибки берется разница между реальным и реконструированным кадром. Так как авторы статьи разрабатывают нейронную сеть для применения в медицине, они столкнулись с некоторыми проблемами, характерными для данной области. Во-первых, с проблемой засорения объектива камеры, что негативно сказывается на точности работы и качестве обучения сети. Во-вторых, при работе с медицинскими данными необходимо учитывать подвижность и эластичность тканей. В результате этого геометрия окружающего пространства может немного изменяться от одного кадра к другому. Третья проблема связана с тем, что биологические жидкости довольно хорошо отражают свет, из-за чего в кадре часто возникают паразитические блики, что также негативно сказывается на точности алгоритма. Для решения этих проблем в процессе обучения к функции потерь был добавлен коэффициент, являющийся маской для изображения. Чтобы исключить вырождение получившейся функции было также добавлено слагаемое, которое штрафует функцию за вырожденную маску. Кроме того, из-за дискретности изображения при построении реконструированного изображения приходится использовать интерполяцию. При больших значениях градиента в окрестности пикселя, интенсивности которого рассчитывается с помощью интерполяции, может возникнуть значительная ошибка. Чтобы это учесть, к функции ошибки было добавлено еще одно слагаемое, для сглаживания описанного эффекта.

В итоге, полученный алгоритм показывает довольно хорошую точность при построении траектории движения камеры. В своей статье авторы приводят сравнение результатов работы полученной нейронной сети с классическими SLAM алгоритмами (Рис. 10), из которого видно значительное преимущество алгоритма, на основе нейронной сети. При работе с медицинскими данными авторы столкнулись с

проблемой низкотекстурируемых областей. Обучаясь на таких данных, нейронная сеть показывала низкую точность при построении карты глубины. Чтобы справиться с этой проблемой, авторы «дообучали» свою сеть на иных, не медицинских данных. В результате качество построенной карты глубины значительно выросло.

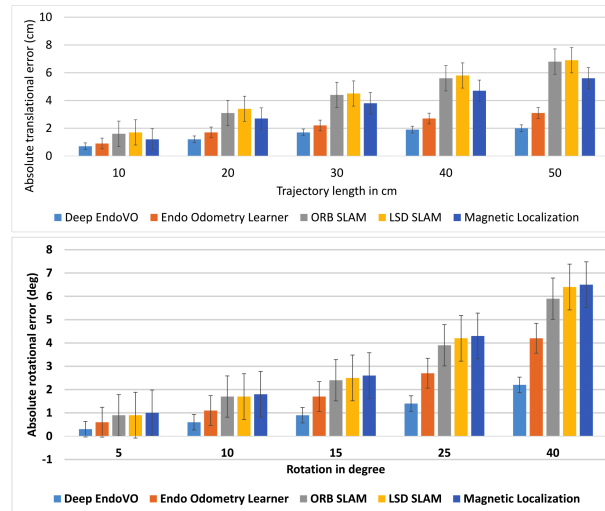


Рис. 10: Сравнение ошибок в определении сдвига (сверху) и поворота (снизу) алгоритма Deep EndoVO и классических SLAM алгоритмов

Суммируя полученные результаты, можно отметить, что использование алгоритмов машинного обучения позволяет добиться лучшего результата в задачах SLAM в условиях малоинвазивной хирургии, чем при использовании классических подходов. Однако, стоит помнить, что алгоритмы машинного обучения требуют большого количества заранее подготовленных данных для обучения, что в рамках малоинвазивной хирургии является проблемой.

Экспериментальные данные

Как уже было сказано, одними из главных компонентов любой работы и любого алгоритма машинного обучения являются экспериментальные данные. Без них невозможно отладить алгоритмы или сравнить их по качеству работы, а также невозможно в принципе создать алгоритмы на основе методов машинного обучения.

Одним из главных датасетов на сегодняшний день является не раз упомянутый датасет центра роботизированной хирургии Хамлина [15], собранный с видеокамер различных роботов, таких как «*da Vinci*» и «*inViva*». Центр Хамлина собрал достаточно большое количество различных датасетов, начиная от просто набора фотографий с монокулярного эндоскопа и заканчивая большими видео со стерео эндоскопа с калибровочной информацией. Однако, не смотря на такое количество данных, в датасетах Хамлина отсутствуют данные об Ground Truth, или иными словами, эталонные данные, по которым в дальнейшем и строится большинство метрик качества.

Другими известными наборами данных являются «*Open-CAS*» датасеты [22]. Среди указанных датасетов можно найти различные наборы изображений и даже трехмерные модели для различных задач сегментации, распознавания и локализации в хирургии. Более того, сайт Open-CAS предоставляет систему валидации полученных данных, что позволяет удобно оценивать результаты работы алгоритма. В случае нашей задачи наиболее интересным является набор Simulation data (синтетические данные). Внутри него представлен большой набор изображений, полученный с виртуального стерео эндоскопа, движущегося внутри виртуальной кишечной полости. За счёт искусственности созданных условий, к изображениям прилагаются данные о Ground Truth и карт глубин для каждой камеры. Однако, к сожалению, виртуальная модель кишечника не способна повторить все условия и сложности реального образца. Так, в виртуальной модели не будет проблем с освещением и отражением, в ней отсутствует жидкая среда и плавающие перед камерой включения. Однако, использовать данный датасет для отладки алгоритма действительно возможно, и как можно было видеть ранее, многие исследовательские группы использовали его для оценки качества своих работ.

Из других датасетов можно отметить множество сборников видео и фото последовательностей, в основном сделанных для задач сегментации и распознавания изображений и операционного поля. Их количество можно оценить как несколько десятков, однако в большинстве своём это скорее небольшие датасеты для собственных нужд исследователей [23, 24, 25, 26, 27]

На этом наборы данных не исчерпываются. Часть исследовательских групп, за неимением лучшего варианта, создают свои датасеты с Ground Truth и реконструкцией трехмерной модели с помощью сторонних датчиков или томографии. Так, представленная ранее команда института интеллектуальных систем Макса Планка [3, 5, 6] использовала магнитные датчики и систему Optitrack для точного позиционирования роботизированного зонда с камерой внутри операционного поля. Это позволило получить им множество данных для обучения и тестирования их алгоритмов. Однако, у этого подхода есть и обратная сторона, особенно если полученные датасеты в дальнейшем не выкладываются в свободный доступ. В таком случае, сравнительный анализ качества работы различных алгоритмов значительно усложняется и чаще всего может быть произведён при условии, что сама команда исследователей решила протестировать множество уже существующих алгоритмов для сравнения.

В итоге можно отметить существование довольно большого числа тестовых данных в свободном доступе, среди которых, к сожалению, существует мало наборов с эталонными данными. С учётом того, что такими являются лишь наборы синтетических данных, можно сделать вывод, что доступных реальных датасетов с эталонными данными не существует. По этой причине некоторые группы исследователей вынуждены создавать наборы данных для своих частных задач, что в итоге несколько усложнило процесс сравнения качества работы алгоритмов.

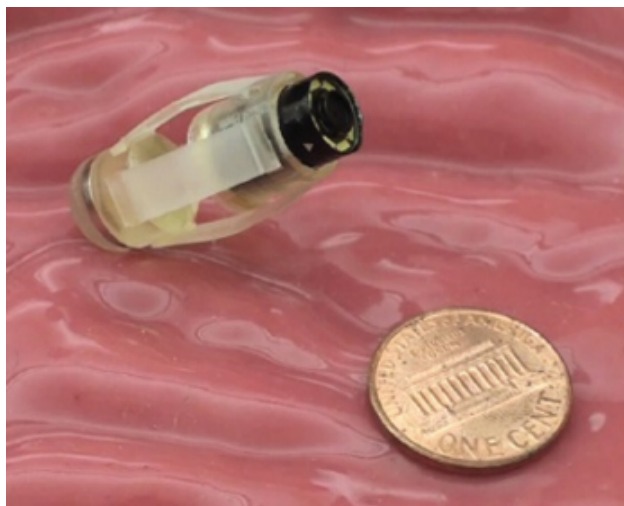


Рис. 11: Фотография зонда, использованного в экспериментах

Заключение

Целью данной работы было исследование состояния развития алгоритмов реконструкции трехмерной модели операционного поля в рамках малоинвазивной хирургии. В частности, мы остановились на наиболее известных и часто используемых семействах алгоритмов SLAM.

До широкого распространения алгоритмов SLAM, задачу реконструкции трехмерной сцены решали в основном алгоритмы SfM. Однако, большого развития в рамках малоинвазивной хирургии они не получили по причине того, что они работают с фиксированным набором изображений, а не с видеорядом, предоставляемым эндоскопами. И всё же некоторые ранние алгоритмы для реконструкции трехмерной сцены использовали SfM метод.

Впоследствии отмечается активное использование алгоритмов семейства SLAM. Алгоритмы данного семейства делятся на 4 типа, из которых особенно выделились 2 — плотные и алгоритмы на основе особых точек. К сожалению, алгоритмы первого типа не особо популярны среди исследовательских команд. И всё же некоторое количество работ было представлено именно по данному типу алгоритмов, в частности, с использованием методов SfS и замыкания.

Среди алгоритмов SLAM, работающих на особых точках, наиболее известным является ORB-SLAM, который и применяется во многих работах. Так, в некоторых работах стараются просто настраивать параметры алгоритма для более оптимальной работы в условиях малоинвазивной хирургии. В других представлены алгоритмы, способные в реальном времени отображать уже обработанную трехмерную модель операционного поля, робастную к деформации различных частей этого поля, и помогающую хирургу лучше ориентироваться во время операции.

В последнее время возросла популярность алгоритмов на основе нейронных сетей. Новые алгоритмы показывают хорошее качество работы и составляют очень серьезную конкуренцию классическим алгоритмам. К минусам таких алгоритмов можно отнести необходимость использования больших вычислительных мощностей для обучения алгоритмов. Кроме того для этих алгоритмов наиболее существенна проблема отсутствия хороших размеченных датасетов.

С наборами данных для задач SLAM малоинвазивной хирургии существуют проблемы - даже при существовании достаточного количества различных наборов изображений и видеозаписей, к сожалению лишь малая часть имеют какие-то дополнительные данные, такие как внутренние параметры камеры или параметры калибровки стереопары. Что же касается данных Ground Truth или восстановленной трехмерной модели, то в открытом доступе существуют лишь синтетические данные, которые при тестировании алгоритма не могут покрыть весь спектр задач, с которыми может столкнуться алгоритм при работе «в поле». Одним из решений проблемы отсутствия датасетов с эталонными данными является создание собственного, что может вызвать большие затруднения у небольших исследовательских команд, так как для создания хорошего датасета необходимо довольно дорогостоящее оборудование.

В заключении можно сказать, что, даже не смотря на некоторые проблемы с экспе-

риментальными данными, алгоритмы трехмерной реконструкции операционного поля и локализации в нём продолжают развиваться и даже начали использовать глубокие нейронные сети. Стоит заметить, что перспективы как использования новых методов, вроде нейронных сетей, так и развития старых, как ORB-SLAM, довольно обширны, и есть множество вариантов для улучшения качества представленных алгоритмов.

Литература

- [1] Колсанов А.В., Каторкин С.Е., Зельтер П.М., Быстров С.А., Колесник И.В., Чаплыгин С.С., Андреев И.С. Виртуальное планирование в абдоминальной хирургии: опыт использования в хирургии печени, селезенки и поджелудочной железы //Клиническая и экспериментальная хирургия. – 2017. – Т. 5. – №. 1. – С. 31-36.
- [2] Stunt J. J. et al. Validation of the ArthroS virtual reality simulator for arthroscopic skills //Knee Surgery, Sports Traumatology, Arthroscopy. – 2015. – Т. 23. – №. 11. – С. 3436-3442.
- [3] Mehmet Turana, Yasin Almalioglu, Helder Araujo, Ender Konukoglu, Metin Sitti. A Non-Rigid Map Fusion-Based RGB-Depth SLAM Method for Endoscopic Capsule Robots //arXiv preprint arXiv: 1705.05444 – 2017.
- [4] Jingwei Song, Jun Wang, Liang Zhao, Shoudong Huang and Gamini Dissanayake. MIS-SLAM: Real-time Large Scale Dense Deformable SLAM System in Minimal Invasive Surgery Based on Heterogeneous Computing //arXiv preprint arXiv: 1803.02009v2. – 2018.
- [5] Mehmet Turan, Yasin Almalioglu, Helder Araujo, Ender Konukoglu, Metin Sitti. Deep endovo: A recurrent convolutional neural network (rcnn) based visual odometry approach for endoscopic capsule robots //Neurocomputing. – 2018. – Т. 275. – С. 1861-1870.
- [6] Mehmet Turan, Evin Pinar Ornek, Nail Ibrahimli, Can Giracoglu, Yasin Almalioglu, Mehmet Fatih Yanik, Metin Sitti. Unsupervised Odometry and Depth Learning for Endoscopic Capsule Robots //arXiv preprint arXiv: 1803.01047 – 2018.
- [7] Hartley R., A. Multiple view geometry in computer vision. – Cambridge university press, 2003.
- [8] Hu M., Penney G., Edwards P., Figl M., Hawkes D.J. 3D reconstruction of internal organ surfaces for minimal invasive surgery //International Conference on Medical Image Computing and Computer-Assisted Intervention. – Springer, Berlin, Heidelberg. – 2007. – С. 68-77.
- [9] Martinec D., Pajdla T. Structure from many perspective images with occlusions //European Conference on Computer Vision. – Springer, Berlin, Heidelberg. – 2002. – С. 355-369.
- [10] Sturm P., Triggs B. A factorization based algorithm for multi-image projective structure and motion //European conference on computer vision. – Springer, Berlin, Heidelberg. – 1996. – С. 709-720.
- [11] Kudan Computer Vision : Different Types of Visual SLAM Systems [электронный ресурс] // <https://www.kudan.eu/kudan-news/different-types-visual-slam-systems/>
- [12] Mur-Artal R., Tardós J. D. Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras //IEEE Transactions on Robotics. – 2017. – Т. 33. – №. 5. – С. 1255-1262.
- [13] Engel J., Schöps T., Cremers D. LSD-SLAM: Large-scale direct monocular SLAM //European Conference on Computer Vision. – Springer, Cham. – 2014. – С. 834-849.
- [14] Long Chen, Wen Tang, Nigel W. John, Tao Ruan Wan, Jian Jun Zhang - Augmented Reality for Depth Cues in Monocular Minimally Invasive Surgery //arXiv preprint arXiv: 1703.01243 – 2017.
- [15] Hamlyn Centre Laparoscopic: Endoscopic Video Datasets [электронный ресурс] // <http://hamlyn.doc.ic.ac.uk/vision/>
- [16] Ping-Sing T., Shah M. Shape from shading using linear approximation //Image and Vision computing. – 1994. – Т. 12. – №. 8. – С. 487-498.
- [17] Habr.com: Что такое свёрточная нейронная сеть [электронный ресурс] // <https://habr.com/post/309508/>
- [18] Habr.com: Как выглядят глубокие нейронные сети и почему они требуют так много памяти [электронный ресурс] // <https://geektimes.com/post/287360/>
- [19] Habr.com: Самое главное о нейронных сетях. Лекция в Яндексе [электронный ресурс] // <https://habr.com/post/309508/>
- [20] Ye M., Johns E., Handa A., Zhang L., Pratt P., Yang G.-Z. Self-Supervised Siamese Learning on Stereo Image Pairs for Depth Estimation in Robotic Surgery //arXiv preprint arXiv: 1705.08260 – 2017.
- [21] Gers F. A., Schmidhuber J., Cummins F. Learning to forget: Continual prediction with LSTM. – 1999.
- [22] Open-CAS [электронный ресурс] // <http://opencas.webarchiv.kit.edu/>
- [23] 3Dircadb [электронный ресурс] // <https://www.ircad.fr/research/3dircadb/?lng=en>
- [24] Surgical Vision@University College London [электронный ресурс] // <http://www0.cs.ucl.ac.uk/staff/M.Allan/#projects>
- [25] Endoscopic Vision Challenge [электронный ресурс] // <https://endovissub2017-workflow.grand-challenge.org/data/>
- [26] Camma Datasets [электронный ресурс] // <http://camma.u-strasbg.fr/datasets>
- [27] HMA Feature-Matching Toolbox [электронный ресурс] // http://ranger.uta.edu/~gianluca/feature_matching

Человеко-знаковые интерфейсы (отчет по итогам годичной работы над созданием «Словаря универсальных по содержанию образов»)*

А. Б. Соломоник

semiosol@netvision.net.il

Протвино, АНО «Институт физико-технической информатики»

*Светлой памяти безвременно ушедшего от нас
Станислава Владимировича Клименко посвящается*

Если мы составляем словарь, то располагаем все помещенные в нем слова по алфавиту соответствующего языка. Это — глобальный принцип словарной работы в любой стране. Он не подходит для собрания образов, которые как и слова пытаются представить и объяснить некие стоящие за образом вещи, состояния либо события. Приходится искать иной способ представления образов в различных их сочетаниях; этому посвящено наше исследование по полученному в РФФИ гранту [1] и настоящая статья. В результате множества попыток и на основании прежнего опыта исследований в этом направлении мы пришли к выводу, что для больших собраний образов организуемым началом может быть только скрытая за образом идея, которая будет одинаково понятна в любой культурной традиции любого народа на земле, хотя она всякий раз будет выражена по-своему.

Ключевые слова: разного рода знаки, слова, образы, идеи за ними стоящие, конструирование собраний образов по скрывающимся за ними сущностными концептами.

Human-Signs Interfaces*

A. B. Solomonik

Protvino, ANO “Institute of computing for physics and technology”

This article is devoted to compilation “dictionaries of images”. Both words and images are kinds of signs demonstrating and explaining something looming beyond them, yet they do it differently. Words in dictionaries are situated according to the alphabet of the corresponding language; whereas images are bereft of this possibility. We found that images of specific qualities may be easily united throughout the idea which conceals behind all of them in the collection (something like Aristo’s $\text{id}\acute{\epsilon}\alpha$). Then these collected images may be understood and utilized by people of various races and cultures with minimum of assistance from gestures and/or simple words, because in spite of different images, ideas lying behind them are concurrent.

Keywords: signs, images, words; collection of images; underlying idea.

Я выделяю три типа реальности: *онтологическую, семиотическую (знаковую) и виртуальную*. Вместе они составляют то, что мы называем нашим окружением или средой. Появляясь на свет, мы имеем дело со всеми этими слоями реальности вместе или по отдельности и чтобы выжить, должны к ним приспособиться. Для этого мы их изучаем, обозначаем и выводим законы, по которым они функционируют. Мы делаем это иначе по отношению к каждому слою реальности, но в данной статье я буду касаться лишь наших отношений с реальностью знакового типа. Поэтому я сосредоточусь

лишь на наших (человеческих) взаимоотношениях с этим видом реальности, или по принятой мной терминологии на *человеко-знаковых интерфейсах*.

Что такое семиотическая реальность и в чем ее основные особенности

Семиотическая реальность появилась позднее онтологической и существенно от нее отличается. Она была создана в ходе познания людьми своего окружения и целиком состоит из знаков, которые являются порождением человеческого ума. Онтология существовала задолго до появления в ней homo sapiens. Думаю, что нет необходимости доказывать обоснованность этого тези-

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 18-07-00225, 18-07-20030.

са; во вселенной мы повсюду сталкиваемся с небесными телами, незаселенными разумными существами. Так же обстояло дело и на Земле. Уже в Святом Писании указано, что вначале был «тóху ва вóху» (сплошной хаос), из которого Бог создал Землю и окружающую ее действительность, затем зверей, птиц и рыб, и в конце — человека. Так что вначале была материальная онтологическая реальность (неживая материя) и лишь потом появились мыслящие существа.

Когда мы познаем окружающую нас онтологическую реальность, мы исследуем ее по созданным людьми правилам и формулируем законы *функционирования тех или иных исследуемых явлений*. Благодаря знанию этих законов мы можем использовать те стороны явлений, которые оказываются нам выгодными, и нейтрализуем те стороны, которые нам вредят. Все формулировки и самих законов бытия и возможностей их использования выражаются придуманными нами же *знаками* (образами, словами, цифрами и пр.). Так создается *знаковая реальность*, которая фиксируется на носителях и передается по всему земному шару, а также будущим поколениям. На этом построены наука, культура и искусство.

Знаковая реальность существует в осязаемом виде (в звуковых записях, книгах, живописи и пр.). Она столь же постижима как и реальность онтологического плана; новые поколения сегодня усваивают законы жизни не только и не столько из непосредственного соприкосновения с онтологией, сколько из книг, радио и телепередач, из устного общения с другими людьми. Успешное обучение жизни зависит от правильного сочетания этих двух источников познания: непосредственной жизненной практики и данных, почерпнутых из знаковой копилки человечества. Но эти два источника, несмотря на то, что они дублируют друг друга, кардинально различаются по своей природе — онтология изначально возникла не по воле человека и развивается в основном по своим заложенным в ней зако-

номерностям; знаковые же источники целиком зависят от тех параметров, которые мы для них сформулировали. Они податливы нашей мысли и могут быть пересмотрены нами в любом направлении, как в правильном, так и в ложном.

Человек полностью господствует в своих взаимоотношениях со знаковой реальностью: он может пользоваться ею сдержанно и опираясь на адекватное отражение соответствующей ей онтологической реальности, а может такую исказить в пользу своих эгоистических соображений. Различия в природе этих типов реальности сказываются на их конкретном воплощении. Знаковая реальность может выдавать такие выводы и заключения, которые не соответствуют законам отражаемой в них объективной действительности.

Все древнегреческие апореи и логические парадоксы, занимавшие людей с незапамятных времен, объясняются частичным несовпадением этих двух типов реальности: и те, которыми оперировал Зенон еще в древние времена; и те, с которыми столкнулись участники первого кругосветного путешествия Магеллана; и те, которые мешали математикам принять учение Г. Кантора о его теории множеств. Путешествие Магеллана доказало, что если обогнуть Землю плывя на Восток, то за год ты сэкономишь календарные сутки, а если поплывешь на Запад, то эти же сутки ты потеряешь, хотя в обоих случаях протяженность пути одинакова. Это происходит потому, что Земля крутится вокруг оси справа налево, и если ты поплывешь против направления ее вращения, ты календарное время сэкономишь, а если ты плывешь по направлению вращения Земли, ты его теряешь. Просто календарь — создание человеческого ума и явление знаковой реальности — не совпадает с объективным ходом времени на поверхности Земли (календарь не может его правильно выразить до конца). То же относится к плоскостным картам, которые не могут правильно отразить круговую поверхность

земного шара, и ко многим иным знаковым произведениям человеческого ума.

Поэтому приходится относиться к объектам знаковой реальности иначе, чем к объектам онтологической реальности, рассматривать их по отдельности и иметь в виду возможные между ними разногласия. Взаимоотношения между человеком и объектами из выдуманных им знаковых систем и будут предметом обсуждения в данной статье, только мы сосредоточимся не на всех знаках, а на их определенной категории — на образах.

Каково место образов в знаковой реальности

Образы — это знаки, которые формируются в нашем сознании по поводу тех явлений или предметов, с которыми мы сталкиваемся. Они могут отражать предметы или явления из онтологии; из знаковой реальности (скажем, из литературных произведений или из живописи); из виртуальной реальности (впечатления от наших снов, например). Сами эти явления не могут быть перенесены в мозг напрямую, но их отражения в мозгу позволяют нам их запоминать, сортировать, модифицировать и пользоваться ими в дальнейшей жизненной практике. Это и есть образы. Они могут быть очень точной копией того, что мы увидели, услышали или почувствовали, а могут быть весьма расплывчатыми; они могут быть детализированы в разной степени, могут быть абсолютно понятными либо оставаться в статусе чего-то неотчетливого и мистического. Они могут оказаться статичными и сохраняться неизменными, а могут быть текучими и динамичными в своей основе. Наконец, они могут быть изолированными, а могут соединяться с другими образами, составляя некую сумму наших впечатлений. Словом, это целая и специфическая категория знаков, которыми мы пользуемся, чтобы перенести наши впечатления извне в наше сознание и там с ними «работать».

Эта категория знаков отличается от иных их категорий. Например, от слов. Словами, как и образами, мы пользуемся, чтобы зашифровать наши жизненные впечатления. Но обе категории знаков коренным образом отличаются друг от друга своими характеристиками и возможностями зашифровать окружающий мир. Если слова могут зашифровать любое наше впечатление, вплоть до самых абстрактных и едва осознаваемых, то образы в значительной мере от них отстают. Эта категория знаков не в состоянии «овеществить» впечатления абстрактного свойства, особенно концепты качественного характера: попробуйте найти точный образ для «иллюзорный» или «лояльный», и вы столкнетесь с непреодолимыми трудностями, в то время как слова «иллюзорный» или «лояльный» легко эти трудности преодолевают. Впрочем, этой проблемой подробно занимались многие философы [2], правда, не совсем в том плане, в котором я подхожу к этой проблеме (см. ниже).

Вообще, любую категорию знаков легко понять, если обратиться к тем средствам, которые восполняют объяснение того или иного знака, выполненное словами. Допустим, мы описываем ту или иную страну. Мы делаем это с помощью словесных знаков; можем сделать очень кратко, а можем — весьма детально. Однако, главными словами во всем описании останутся название страны и краткое ее описание (когда возникла, что в ней происходило, что она представляет собой сейчас). Словесное описание может быть проиллюстрировано фотографиями (образами), а в этнографическом музее и непосредственными артефактами, вывезенными из этой страны. Наконец, мы можем оперировать цифрами, что даст нам представление о пространственных размерах страны, ее населении, ее промышленности, сельском хозяйстве и пр. Вы тут же почувствуете специфику каждой категории знаков, использованных для объяснения.

Образы чаще всего необходимы для получения внешнего облика описываемого яв-

ния. Мы можем написать целую книгу о Шопенгауэре, но не получить о нем полного представления без его фотографии или портрета. С другой стороны, мы не сможем понять, что представляет та или скульптура в парке без ее словесной атрибуции. Эти отношения между образами и словами нам пригодились для выполнения проекта по созданию «Универсального словаря образов». Ни один образ в нем не работает без фотографий или схематичного изображения того или иного объекта, с одной стороны, и без его словесного пояснения, с другой.

Есть еще одно отличие образа от словесного знака. Словесный знак всегда включает полную картину своего референта. Допустим, слово «человек» обозначает всего человека без остатка, в то время как фотография человека изображает его лишь в определенном ракурсе. Для полного представления человека во всем его объеме требуется множество фотографий одного и того же человека с различных точек зрения. Поэтому в словаре образов какой-то объект онтологической действительности должен быть повторен несколько раз с разных позиций и под разными углами зрения. Число таких отображений может быть бесконечным, поэтому мы довольствовались лишь некоторым числом образов того же объекта, отражающих по нашему мнению его главные особенности. В каждом конкретном случае требуется учитывать и доступное пространство, предоставляемое нам для нужного образа, а это резко влияет на его размеры и на число дополняющих друг друга возможных вариантов изображения.

По этому параметру (объем изображаемого в объекте) слово явно превышает возможности образа. Да и по глубине изображаемого оно тоже многогранней. Слово «человек» ставит объект в соседство и в сравнение со всеми аналогичными предметами и явлениями — с живыми и неживыми предметами, а как индивидуума — с иными членами социума и пр. Мы говорим: «Вот это человек!», имея в виду нечто особенное. Фо-

тография человека не имеет таких богатых связей; она просто образ и не более того. Поэтому для законченного представления объектов из онтологической действительности всегда требуется некоторое сочетание образов и слов — они дополняют друг друга: слова восполняют недостаточность образов для конкретной характеристики его качеств, а образы — для дополнения словесных характеристик конкретными представлениями объектов.

Зато в представлениях *виртуальной реальности* образы играют куда более значительную роль, нежели слова. Слова лишь называют предметы и явления виртуальной реальности, а образы представляют их в самых живописных деталях. Образам доступны любые фантастические задумки, в то время как слова затрудняются их определить. В самом деле, обычно таких существ в реальном мире нет — и мы открываем шлюзы нашего воображения, чтобы их представить. Никакими реальными границами при этом мы не ограничены, представляй, что хочешь. После появления компьютеров в их приложениях появились многочисленные образы самого невероятного плана, особенно в компьютерных играх, и этот жанр пользуется заслуженным успехом у пользователей младшего возраста, да и у людей постарше. Все же люди более зрелого поколения привыкли себя сдерживать образами, почерпнутыми из реальной жизни; для незрелого же ума это — не помеха, а, наоборот, поощрение.

Еще одно качество образов

К упомянутому выше качеству образов примыкает еще одно их свойство, может быть, самое главное для нашего их понимания. Обладая способностью овеществлять воображаемые предметы и их характеристики, образы дают возможность аллегорического воплощения наших обычно скрытых, но столь ценимых человечеством устремлений — к добру, справедливости, щедрости, благородству и иным благостным (как, впрочем, и не благостным) иде-

алам. Их, разумеется, можно обозначать просто словами; но куда таким обозначениям до аллегорических выражений этих качеств в живописи или скульптуре, на театре либо в кино. Самое важное для нас заключается в том, что с помощью такого рода обращений к какой-то идее (своего рода аристотелевской «форме»), мы можем отбирать целые серии образов, эту форму воплощающих. Чем мы и пользуемся, собирая и ранжируя конкретные образы по определенной теме в нашем проекте.

Образов так много, что их отбор возможен только после установления скрытой идеи, их объединяющей. Никакого иного способа не существует. Для слов, которых тоже неизмеримое количество, нашли иной порядок их описания — по алфавиту, который, совершенно не касаясь содержания перечисляемых слов, расставляет их по легко обнаруживаемым местам. Таким способом создаются *общие для данного языка толковые словари*, куда входит лишь небольшое число специальных терминов. Если мы хотим создать словарь специализированных терминов (по отраслям знания, по какой-то профессии), мы отбираем туда соответствующие термины и располагаем их опять таки в алфавитном порядке. Такого рода словари называются *тезаурусами*. Иногда в тезаурусах практикуются подразделения, принятые в данной науке либо профессии. Так, в химии мы можем ввести два основных подразделения: отдельно для терминов *органической* и отдельно для терминов *неорганической* химии.

Алфавитный принцип невозможен для словарей образов. Что же возможно? Возможно отбирать образы только согласно скрытой за ними и объединяющей их в едином комплексе идее. При этом мы как бы выбираем какой-то стержень, проявляющийся в жизни любых людей, где бы они не проживали, и по нему отбираем наличествующие в нем характеристики данного конкретного проявления. В этом и сказывается *универсальность отбираемых в словаре образов*. Они будут различными для разных куль-

тур и обычаев, но объединяющая их вместе суть, сущность будет *универсальной*. Скажем, формы проявления уважения к старым людям в разных социумах абсолютно разнообразны, но лежащая за ними идея одна и та же — почтение к старым и слабым. Или — всем людям на земле надо есть. Прием пищи в разных культурах отличается по форме, но не по содержанию. Поэтому образы, отражающие это явление в разных его проявлениях, будут *универсальны* по лежащей за всеми ними идее.

На это обстоятельство обращали внимание многие исследователи. Так М. Я. Розенфельд пишет в монографии «Слово и образ»: «Представление любого человека о конкретном предмете, называемом словом, нам действительно неизвестно — оно индивидуально; но в индивидуальном представлении любого человека есть часть, общая для всех говорящих, и эта часть *наиндивидуальна*» [3]. Эту самую *наиндивидуальность*, которую я определяю, как аристотелевскую *форму*, мы и можем использовать для выделения некоторой группы образов, собранных вместе для характеристики данной универсальной человеческой идеи. Идея должна реально проявляться в жизни любого человека в любом социуме, и только в этом случае она оказывается *универсальной*.

Это кардинальное для нашего проекта положение было мной понято лишь в ходе непосредственной работы над словарем в течение этого года. И в результате я изменил прежнее название проекта — *Универсальный словарь образов* — на *Словарь универсальных по содержанию образов (Pictorial Dictionaries of Universal Ideas)*. В действительности нет одинаковых универсальных образов, они в разных условиях приобретают специфические особенности. Зато есть универсальные идеи, по поводу которых появляются многочисленные образы, что, по-моему, и отражается во втором наименовании проекта.

На пути универсальности встает то обстоятельство, что в каждом отдельном социуме

та же самая идея обретает свой специфический вид в соответствии с местными традициями и обычаями. Где-то *идея приветствия* реализуется в рукопожатии, а где-то в сложенных перед грудью руках. У одного народа понятие о гостеприимстве выражается в виде предоставления случайно пришедшему путнику пристанища, а у другого — предоставление ему еще и своей жены. На пути к глобальному обществу встают самые разнообразные проявления одного и того же явления — их надо изучить и представить людям, которые их не знают, и из-за этого могут столкнуться с непредвиденными обстоятельствами. Я рассматриваю мой проект как один из шагов в этом направлении. Если я покажу в образах некое явление, а кто-то еще прибавит по тому же поводу иной обычай, а третий вставит в возникающий ряд образов еще один вариант, можно будет предотвратить возможные неприятности такого рода. В этом и смысл нашего проекта.

Детали воплощения нового понимания проекта

После появления новой направляющей идеи все встало на свои места. Для пробы мы попытались составить небольшой словарь (примерно на 100 страниц) для туристов в Москве. Мы вычленили основные категории образов, которые необходимо представить туристам (иностранным и российским) для ориентации в географических пределах Москвы, для обеспечения себя всем необходимым и удовлетворения материальных и духовных потребностей. Могут возразить: «А что тут нового, будто не было и раньше путеводителей для туристов в Москве?». Дело в том, что мы создаем туристический словарь образов не только для его практического использования, но, прежде всего, для доказательства тех теоретических установок, которые мы для себя поставили: «*Вот универсальная идея — прибытие в чужой город по делам либо ради отдыха. Что важно включить в объем информации для людей, которые в данном*

месте не были и не знают принятого в ней языка; и сделать это в виде образов, а не в виде, скажем, словесных пояснений, которые нами используются лишь для раскрытия имеющихся в образе деталей?» Ниже приведены разработанные нами категории образов, которые мы включили в словарь:

1. Московские аэропорты, железнодорожные вокзалы и внутригородской транспорт.
2. Ориентация на просторах города (обозначение улиц, площадей. . . , нумерация домов, подъездов и квартир).
3. Переход улицы по семафору и без.
4. Обозначения различных объектов (аптек, больниц, магазинов, театров, кинотеатров, общественных туалетов. . .).
5. Ориентация внутри крупных зданий (табло с обозначениями, места для курения, кафе, туалеты и пр.)
6. Объекты религиозного назначения (церкви, мечети и синагоги).
7. Отели и хостелы, места туристической информации.
8. Деньги (банкноты и монеты).
9. Наиболее популярные театры и кинотеатры, консерватория, планетарий, дельфинарий.
10. Спортивные сооружения (стадионы, фитнес клубы, бассейны).
11. Места для рекреации (парки, сады, зоосады, дома культуры).
12. Символы Москвы (Кремль, Александровский сад, Ленинская библиотека, памятники Пушкину и Долгорукову).
13. Принятые жесты: встреч и прощаний, согласия и несогласия, жесты успеха и поощрения и некоторые другие.
14. Неодобряемое поведение в образах (переход улиц в неположенном месте, распитие напитков в садах и на улицах, бросание окурков на тротуар, оставление тряпок, бумажек и бутылок не в помойках).

Даже беглого взгляда на этот список достаточно, чтобы понять, что это не обычный перечень туристических объектов, но стремление раскрыть сущность идеи при-

бытия в незнакомое место, язык которого тебе неизвестен, но где тебе придется провести некоторое время и сделать это с удовольствием и пользой для себя. Туристическое содержание словаря выбрано для примера, на котором легко показать сущность отображаемого в образах и продемонстрировать хорошо известные составителям детали иллюстрируемого явления. Среди деталей включены и такие, которые отсутствуют в обычных туристических справочниках, например, правила представления названий улиц и нумерации домов, ознакомление с принятыми в обиходе жестами и предупреждения о нежелательном поведении. Словарик наш еще не сверстан до конца, поэтому я приведу примеры только из самого начала списка, см. рис. 1, 2.

В ходе работы по составлению словаря выкристаллизовывается новая терминология, касающаяся как технологии производства, так и содержания отображенных в словаре вещей. Имеется в виду собрать в экспериментальном словаре около ста **фреймов**. Фреймом мы назвали законченный сюжет словаря, своего рода словарную статью, например, «Шереметьевский аэропорт под Москвой». Фрейм может включать одно либо несколько воплощений, каждый из которых называется **кадром**. «Схема Шереметьевского аэропорта» конкретизируется еще тремя кадрами, изображающими три из имеющихся в порту терминалов. Дополнительно аэропорт включен в схему связей всех аэропортов с Москвой. Так что фрейм «Шереметьево» фактически реализуется пятью кадрами, а количество образов в них много больше (столько, сколько мы считали нужным показать для ориентировки пассажиров, прилетающих либо улетающих из аэропорта).

В кадр может быть включено много **образов (имиджей)**. Сам образ может быть **одиночным** либо **продолженным** (жест согласия либо несогласия, выраженный на экране длительным покачиванием головой). Образ также может быть включен в **мизансцену** (вход, движение, спуск по эс-

калатору и выход на платформу в метро — см. ниже на рис. 2). Каждый образ сопровождается **объясняющими надписями** на русском и на английском языках. По исполнению словарь будет включать два типа образов: создание **коллажей**, изготавливаемых на столе дизайнера, и фотографирование на местности (в Москве) в сопровождении с актерами, разыгрывающими сценки либо без них. Эти сценки в виде кадров потом включаются в словарь.

Наконец (и это весьма важное замечание), в результате мы получаем законченный продукт нового вида лексикографии — лексикографии образов (имиджей). Я предлагаю назвать такую лексикографию **идеографической**. Производными от этого термина будет прилагательное **идеографический** и существительное **идеограф** — человек, занимающийся созданием словарей образов. Сам термин **идеографика (идеография)** не абсолютно нов для обозначения особого периода в истории человеческого письма. Он использовался как раз для образных изображений на письме: «Идеографическое письмо (от греческого *idea* — идея, образ и *grapho* — пишу), принцип письма, использующий **идеограммы** — схематизированные рисунки, передающие понятия, ассоциируемые с изображениями. Идеографическое письмо не воспроизводило связную речь. Оно предшествовало собственно письменности. К идеографическому письму восходит словесно-слоговое письмо. Идеографический характер имели древнеегипетская, шумерская и другие старейшие системы письма. Наибольшего развития достигло в китайской иероглифике» [4].

Имея в виду изложенные выше соображения о том, что образы для конкретного словаря должны отбираться на основе их принадлежности к тому или иному скрывающемуся за ними общему понятию или концепту, термин **идеографика** очень точно обозначает тот вид лексикографии, к которому относится наш проект. Многочисленные попытки создания словарей **универсальных образов** заканчивались неудач-



Рис. 1: Схема Савеловского вокзала

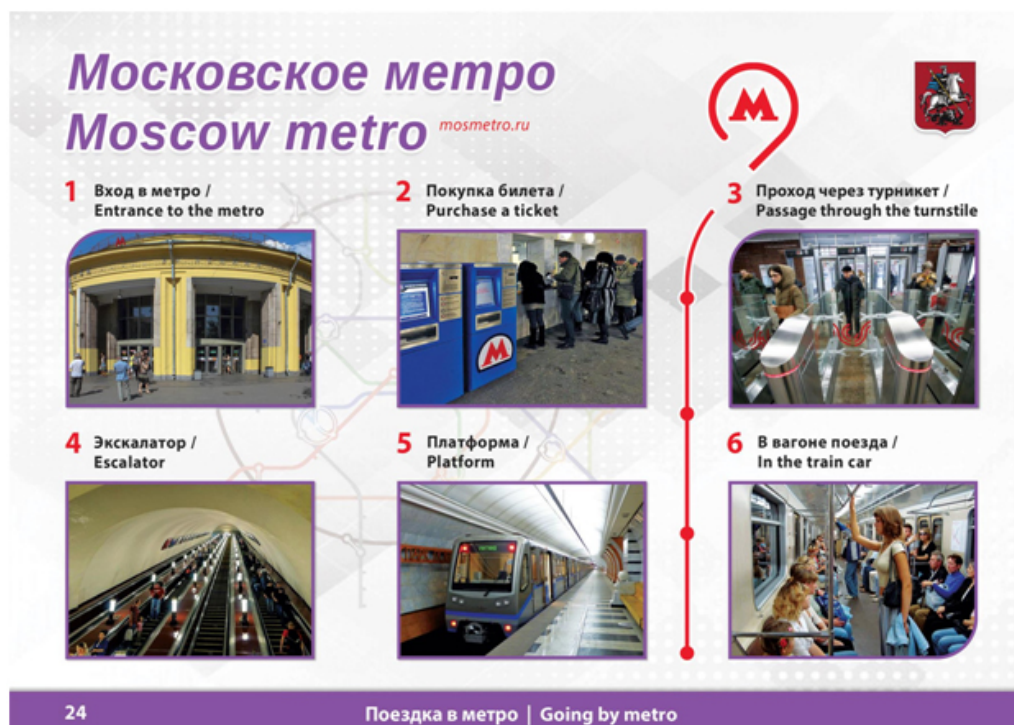


Рис. 2: Поездка в метро

но, потому что всеобщих образов пока что очень мало. Зато есть идеи универсального плана, которые можно выразить с помощью образов, чему и посвящен наш экспериментальный словарик. Мы надеемся, что после его завершения к нам присоединятся многочисленные желающие выразить образами свои собственные идеи аналогичного плана. Тогда возникнут подобные нашему словари для туристов в иных местах земного шара — как в самой России, так и за ее пределами.

Если мы, действительно, получим широкую поддержку и наберется достаточно много фреймов по той или иной теме из нескольких стран, можно будет ходатайствовать перед ЮНЕСКО о создании комиссии, которая отберет стандартные образы, отражающие конкретную идею однообразно по всему миру. После этого надо будет выбрать другую концептуальную идею и продемонстрировать ее в образах. Собрание всех возможных идей будет представлять *идеографику* в глобальном масштабе на каждый данный период времени.

Инициирование дубляжа наших заготовок в различных странах и регионах на местном материале является одним из приоритетов проекта. Это было заявлено в первоначальной заявке на проект и будет нашей задачей в будущем, если работа будет продолжена. Тогда возникнет много центров создания словарей подобного рода и использование концептов *идеографика*, *идеографический*, *идеограф* (равно как и иных упомянутых выше терминов) станет привычным. Возникнет также много больше, чем сегодня, принятых повсюду одинаковых образов. Это — одно из фундаментальных намерений нашего исследования. Это также даст новый импульс науке, которая в российской номенклатуре называется *обществоведением*.

Литература

- [1] Научный проект №18-07-00225 «Универсальный словарь образов», получил грант от Российского Фонда Фундаментальных Исследований по результатам конкурса 2018 года.
- [2] Шопенгауэр Артур. Мир как воля и представление. Москва, «Эксмо», 2015.
- [3] Стернин И.А. и Розенфельд М.Я. Слово и образ. Воронеж: «Истоки», 2008, с. 13.
- [4] https://illustrated_dictionary.academic.ru (апрель 2018).

Когнитивная информатика и гуманитарные технологии*

V. V. Rykov

vlad.rykov@gmail.com

Московский физико-технический институт (Государственный университет)

Информационные процессы происходят в мире знаков. В них так или иначе соучаствуют люди. Это значит, что когнитивные модели, которые в них реализованы, так или иначе базируются на семиотических т. е. гуманитарных технологиях или универсалиях. В последние годы, особенно с появлением соцсетей, все большую значимость начинают играть модели и универсалии, основанные на науках о коммуникациях, и прежде всего на древнейшей из них — риторике. Проблемы применимости этих моделей к информационной практике обсуждаются в данной статье.

Ключевые слова: когнитивные технологии, универсалии, коммуникация, семиотика, риторика.

Cognitive Informatics and Humanitarian Technologies*

V. V. Rykov

Moscow Institute of Physics and Technology (State University)

This paper describes the crucial approaches in cognitive technologies. The info processing runs in the world of signs. It means that they are based on semiotic universals. They are often based on communication. So the rhetorical universals are available to them. So this paper discusses the problems of humanitarian models application to info processing. The important feature of proposed technologies is its availability for an engineer.

Keywords: cognitive technology, models, universals, rhetorics, communication, semiotics.

В процессе своей жизнедеятельности при взаимодействии с различными объектами реального окружающего мира, а также идеальными объектами и объектами знаковой природы, люди придумали много приемов, технологий, теорий и инструментов как оптимизировать эту деятельность. С середины XX века начала бурно развиваться информатика, которая резко повысила эффективность работы со знаковыми объектами. В последние годы все более часто стали применяться когнитивные модели и технологии, которые имеют дело с так называемыми когнитивными объектами или обрабатывают знание.

Знание имеет качественно другую природу по сравнению с информацией и требует построения принципиально новых систем его обработки. Элементарная система обработки знаний включает в себя как минимум две части — базу знаний (knowledge base — KB) и когнитивный процессор (cognitive processor — CP), который ее обрабатывает. Результатом их совместной работы является знание, которого не существовало до сих

пор и которое предназначено оптимизировать ту или иную сторону производственной деятельности. Очень часто составной частью когнитивной системы (КС) является человек, который не только запрашивает новое знание в соответствии со сложившейся ситуацией для дальнейшего использования, но и сам соучаствует в процессе порождения нового знания. Когнитивный эффект применения знания лучше всего продемонстрировать на некоторых существующих моделях [1, 2, 3, 4].

Важно сразу отметить следующее. Для более полного и адекватного понимания логики работы описываемых моделей следует иметь в виду, что они базируются на многих традиционных и давно существующих в гуманитарных науках — таких как семиотика (наука и знаках), теориях взаимодействия текста и человека, а также теории эффективной знаковой деятельности, базирующейся на постулатах риторики. Природа знания и возможности его обработки всегда активно разрабатывалась различными философами. А алгоритмы порождения нового знания были разработаны еще Аристотелем (силлогизмы). Увидеть эти модели может только тот, кто знаком с ними. Этот

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20030.

первый используемый здесь когнитивный эффект называется видение, управляемое знанием (knowledge based vision — KBV).

Эффективность этих моделей, которые зачастую неявно использовались и ранее, можно продемонстрировать на примере информационной поддержки, принесшей победу на выборах Дональду Трампу. Рассылка избирателям была избирательной и базировалась на классической риторической модели образа автора [3, 4, 5].

Прежде всего требуется увидеть и осознать эти когнитивные модели (КМ). Начнем с модели, которая называется Бритва Оккама. Она гласит, что из многих моделей, объясняющих что-либо, истинной является самая простая. На этом принципе мы будем основываться в дальнейшем.

Следует в самом начале определить еще одну базовую метамоделю — в какой форме формализуются или выглядят закономерности в мире информатики? — Ответ — в форме универсалий. Универсалия — это значит, что эта закономерность допускает исключения, оставаясь при этом верной. Это так и есть, но не всегда это ясно осознают. Показательный пример — универсалия — если прыгнуть из самолета с парашютом и он не раскроется, то парашютисту грозит смерть. Вопрос — опровергают ли эту универсалию достоверные сообщения о том, что каждый год десятки людей в таких случаях избегают смерти? — Те, кто считает, что эта универсалия неправильная, будут ли прыгать без парашюта? Более сложные и плодотворные примеры универсалий будут приведены позже. А сейчас нам достаточно осознавать то, что мы живем в мире универсалий. Только в таких науках, как физика или математика появление исключений вызывает революцию. Так, например, обнаружение фотоэффекта в конце XIX века вызвало революцию в классической физике и появление квантовой механики и других подобных наук. Это не значит, что в информатике универсалии не корректируются. Но отдельные исключения там не смертельны.

Другая базовая КМ — это системный подход, который часто сводится к определению, что такое система и следующему за этим каскаду формул. Однако очень часто он эффективно работает и без формул. Достаточно рассмотреть — частью какой системы является проблемный объект. Например, никогда мы не поймем, почему кукушка откладывает свои яйца в чужие гнезда, как бы мы ни изучали ее биологию и повадки. А решение появляется, когда мы рассмотрим кукушку как часть системы, в которой она живет — а именно леса. А в лесу есть гусеницы с настолько острыми колючками, что ими невозможно кормить птенцов. Несложными системными рассуждениями нетрудно прийти к выводу о том, что лесу как системе пришлось выделить одну птичку для поддержания баланса, которая оказалась бы способна поедать таких гусениц. А ее птенцов взять на воспитание для сохранения ее рода. Рассуждение несложное, но нетривиальное. А еще более нетривиальным выглядит такой факт. В те годы, когда таких гусениц мало, и не нужно так много кукушек — лесные птички внезапно прозревают и выкидывают из своих гнезд кукушат! Моделей, объясняющих этот эффект еще нет и не предвидится. Это естественным образом ограничивает пределы наших когнитивных возможностей.

Другой пример системного подхода также может быть показательным. По ТВ и другим источникам можно часто услышать про Стоунхендж. Этот самый большой кромлех — а это называется на Британских островах именно кромлех — и их развалины до сих пор можно много увидеть на полях Англии и Ирландии. Утверждается (заметим — авторы — британцы — об этом есть фильм BBC), что Стоунхендж (заметим — про другие кромлехи не говорят) построили предки британцев несколько тысяч лет назад, так как они очень интересовались астрономией. Понятно — в отличие от более примитивных жителей континентальной Европы. Попробуем применить и здесь наш системный подход.

То есть здесь мы снова применим наши две метамодели — Бритву Оккама — как поиск самого простого объяснения, применяя при этом еще одну метамодель — KBV (Knowledge Based Vision). Системный вопрос — в каком мире, в какой климатической системе жили доисторические но уже земледельцы обитатели Британских островов? Ответ — эти острова были и есть на широте Киева. И сверху на них падало примерно то же что и на древних и современных киевлян. Но на британской земле из-за теплого Гольфстрима творилось совсем другое. Снег таял в течение суток. Дожди и снега из-за близости океана шли почти непрерывно день и ночь. Не зря даже в современной Британии до сих пор основная тема бесед — о погоде. Правда — в теплых салонах! Все, о чем молили богов древние местные жители, это знать когда наступит весна. Не вовремя посеянное давало худший урожай или вообще не всходило. Эту единственную дату — время весеннего солнцестояния им показывали кромлехи, которые были построены в каждой местности. Это еще раз проверили американские информатики, прокрутив координаты всех камней Стоунхенджа на компьютере [1].

Как мы видим, наиболее сложным и нетривиальным системный подход выглядит там, где нет формул! С учетом того, что обнаруженная закономерность есть универсалия и является истинной согласно Бритве Оккама. Ничто не мешает проводить вокруг Стоунхенджа различные ритуалы, начиная со Средних веков, а может и раньше. Еще раз. Мы увидели разгадку кромлехов как универсалию, истинную в рамках Бритвы Оккама, которая никогда не была загадкой для местных жителей, используя для этого системный подход в рамках модели видения, управляемого знанием (KBV). Такие результаты дает системный подход. Однако в современной информатике остается место и для других когнитивных моделей (КМ). Начнем с простого. В каком мире происходят информационные процессы?

Долгое время этот вопрос не был актуален. Информатики решали более практические задачи. Начнем с простого, как учит нас Оккам. Посмотрим на школьника, играющего в компьютерную игру, вместо того, чтобы делать уроки. Где происходят события этой игры? Классическая философия учит нас, что есть только два мира — материальный и идеальный. — События игры, конечно же происходят не во внешнем мире и тем более не в голове школьника — иначе ему не нужен был бы компьютер. Значит, они происходят в третьем мире — мире знаков, материализующихся, правда, на экране компьютера. — Примерно, как и словесные знаки и формулы, материализующиеся на бумаге книг, журналов или на учебной доске. Феномен третьего, знакового мира был открыт Григорием Сковородой в конце XVIII века — во времена Екатерины. Он использовал для открытия этого мира конечно не компьютерную игру, а Библию. Мир Библии — в мире знаков. Но материализован он в виде толстой книги. И опубликовано это открытие — что важно — в одном из его трактатов. Но это открытие не было востребовано в те времена. Тогда востребовано было другое. Востребовано это было в середине XX века и переоткрыто американскими информатиками. Они так его и называли — третий мир [1]. Однако — одна из универсалий научного мира гласит — первенство в открытии за тем, у кого первая публикация. И здесь мы видим как действуют исключения из этого правила. Никто не упоминает Сковороду. Но открыл мир знаков он.

Итак, информационные процессы (ИП) происходят в мире знаков. Действительно, вывод тривиален. Компьютеры обрабатывают не идеи и не материальные объекты, а их знаки. Семиотика, наука о знаках, традиционно делится на синтактику, семантику и прагматику [1]. И здесь можно обнаружить много полезных универсалий. И далее, на их основе, сможем по-другому увидеть многие ИП и сделать нетривиальные выводы. Рассмотрим для примера машин-

ный перевод (МП). Пока что все системы МП строятся на базе синтактики. То есть они не знают семантики текстов, которые переводят. Рассмотрим английское предложение — I put my table on his table. Слово table имеет как минимум два перевода — стол и таблица. Современной системе МП ничего не стоит разобрать синтактику этого предложения. Но какое из двух значений ставить в перевод? Получается четыре варианта. А сколько их будет после перевода последующих предложений? Тем не менее, системы МП действуют. Не всегда совершенно, но переводят. Здесь мы видим еще одну важную универсалию — если очень нужно, то можно. Шуточный вариант — если нельзя, но очень хочется — то можно. Действительно, о МП начали мечтать еще в 60-е годы XX века, думая, что достаточно одолеть синтактику. Было придумано много усовершенствований, позволивших привести МП к приемлемому виду [1].

Обратимся к наиболее казалось бы гуманитарной науке — риторике. Ее законы были сформулированы еще более 2000 лет назад Аристотелем. Какие ее универсалии или технологии могли бы пригодиться в информатике? Начнем с простого — образа автора. Это совершенно не то, что есть автор. В коммуникации, а это мир знаков, люди общаются через свои образы, профили, аватары и т. д. И сейчас, особенно в соцсетях, можно наблюдать интересные закономерности о том, как ведут себя люди в зависимости от своего профиля. И наоборот, конструирование профиля человека по его текстам, лайкам и т. п. в Интернете оказалось важной задачей. Более того, технология рассылки различных агиттекстов, основанная на различиях в оцифрованных моделях образа автора избирателей, принесла победу Дональду Трампу на выборах в 2016 году [5].

Обратимся к другой важной риторической категории — теории общих мест. Это некоторые общепризнанные семантические универсалии, которые должны быть отражены в том или ином тексте. Например, в

каждом дипломе или диссертации должны быть описаны уникальные для данной работы, но общепризнанные общие места — такие как новизна, актуальность, практическая значимость. Теория общих мест позволяет нам строить гораздо более значимые и интересные результаты — например, алгоритмически порождать уникальные изобретения [6].

На основе вышеизложенного мы можем прийти к следующим выводам:

1. Информационные процессы (ИП) происходят в мире знаков. Это задает определенные особенности и накладывает определенные требования на эти процессы а также помогает лучше их понимать.
2. Строгие законы логики, математики, физики имеют непреходящее и неоспоримое значение. Но в знаковом мире, где реализуются ИП, действуют модели, имеющие свои отличительные особенности. Эти ИП реализуются в мире универсалий, частью которых зачастую является человек, и которые допускают исключения, не влияющие на их истинность.
3. В мире ИП очень эффективным и продуктивным является системный подход. Очень важно бывает понять и реализовать в форме каких универсалий и других формализмов реализуется данный ИП, который как правило является сам по себе сложной системой и является частью не менее сложной системы.
4. К проектированию, реализации и пониманию многих ИП применимы универсалии классической риторики — такие как образ автора (участника ИП), а также методика изобретения по общим местам.
5. Для понимания и проектирования ИП, происходящих в знаковом мире, следует сознательно применять законы семиотики — науки о знаках. То есть понимать и уметь формализовать их синтактику (состав, взаимосвязи — т. е. систему), семантику — значимость и функции их ча-

стей и целого ИП — и прагматику — доминанту и особенности их функционирования.

Литература

- [1] Рыков В.В. Обработка нечисловой информации. Управление знаниями. — М.: МФТИ, 2007.
- [2] Рыков В.В. Сервисные модели инновационной деятельности // Труды 59-й научной конференции МФТИ. — М.: МФТИ, 2016.
- [3] Рыков В.В., Колчина В.А. Реконструкция образа автора как когнитивный процесс // Труды 59-й научной конференции МФТИ. — М.: МФТИ, 2016.
- [4] Рыков В.В. Образ автора - автоматическое формирование // Проблемы риторики, поэтики и стилистики в современном информационном обществе. — Астрахань: Издательский дом «Астраханский университет», 2016. — С. 41-45.
- [5] Антонова Т. Наск уоц, миссис Клинтон! // Московский комсомолец. — М., 2017. — N. 28.
- [6] Альтшулер Г.С. Алгоритм изобретения. — М., 1969.

Мультимедийные коммуникации — критерии классификации*

М. С. Белова, В. В. Рыков

belova.margarita11@gmail.com | vlad.rykov@gmail.com

Московский физико-технический институт (Государственный университет)

Информационные процессы происходят в мире знаков. В них так или иначе соучаствуют люди. Это значит, что когнитивные модели, которые в них реализованы, так или иначе базируются на семиотических т. е. гуманитарных технологиях или универсалиях. Проблема коммуникаций, в которых соучаствуют жесты, или — в более широком смысле мультимедийных коммуникаций приобретает большое значение. Прежде всего важно — является ли жест или другой объект частью коммуникации, какова его функция, а также классы коммуникативных объектов.

Ключевые слова: слова: когнитивные технологии, универсалии, коммуникация, семиотика, жест, мультимедийные коммуникации, классификация.

Cognitive Informatics and Humanitarian Technologies*

M. S. Belova, V. V. Rykov

Moscow Institute of Physics and Technology (State University)

This paper describes the crucial approaches in cognitive technologies. The info processing runs in the world of signs. It means that they are based on semiotic universals. They are often based on communication.

So the regular universals are available to them. And this paper discusses the problems of humanitarian models classification based on gestures and other signs and its application to info processing. The important feature of proposed technologies and classifications is their availability for an engineer and other applications.

Keywords: cognitive technology, models, universals, gesture, classification, multimedia communication, semiotics.

В процессе своей жизнедеятельности при взаимодействии с различными объектами реального окружающего мира, а также идеальными объектами и объектами знаковой природы, люди придумали много приемов, технологий, теорий и инструментов как оптимизировать эту деятельность. С середины XX века начала бурно развиваться информатика, которая резко повысила эффективность работы со знаковыми объектами. В последние годы все более часто стали применяться когнитивные модели и технологии, которые имеют дело с так называемыми когнитивными объектами или обрабатывают знание [1, 2, 3].

Коммуникация часто сопровождается специфическими семиотическими объектами — жестами и другими знаками. Можно сказать иначе — жесты являются часто неотъемлемой частью коммуникации. Тема исследования мультимедийной коммуникации и, в частности, жестикуляции становится все более актуальной.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20030.

Прежде всего надо определить конкретно жест как объект наиболее частотный и наглядный как объект научного исследования. Адам Кендон определил те свойства жеста, которые отличают жест от других движений человеческого тела [5]. Это пять свойств жеста, которые отличают его от обычных незначительных движений человеческого тела. Это — осознанность, намеренность, информативность, контролируемость и бескорыстие жеста. Эта классификация не всегда однозначна, хотя и полезна. Так, осознанность жеста не всегда бывает. Часто жесты бывают бессознательны. Аналогично с намеренностью. Часто люди жестикулируют и ненамеренно. Это касается и информативности жеста. Помахивание рукой в качестве приветствия может быть просто обычным взмахом. Это относится и к другим упомянутым выше качествам, что не отрицает их полезности, т. к. они являются универсалиями — т. е. когнитивными законами, допускающими исключения. В качестве примера можно напомнить, что универсалия, гласящая, что если парашют не раскроется, то парашютист разобьется,

не перестает быть верной от того, что каждый год десятки парашютистов при этом остаются живы.

Например, кивок головы может быть: а) частью движения человека при ходьбе, б) молчаливым выражением согласия и в) сопровождать словесное выражение согласия. Эти и не только эти факторы определяют трудности при классификации лингвистических жестов.

Для решения этих проблем были выдвинуты соответствующие теории. Так, лингвист Кендон предложил различать жесты по таким трем показателям: 1) обязательность словесного сопровождения — т. е. могут ли жесты реализовываться без речи, 2) наличие или отсутствие системных языковых показателей при жестикуляции и 3) степень повторяемости или регулярности жеста — воспроизводит ли его говорящий постоянно и регулярно или случайным образом [5].

Систематизируя классификацию Кендона, другой лингвист Дэвид Макнилл позиционировал типы жестов, описанных в работах Кендона в виде некоторой координатной прямой, которую он назвал «Континуум Кендона»:

Gesticulation — Language-Like Gestures — Pantomimes — Emblems — Sign Languages [6].

Слева направо: 1) обязательность речи, сопровождающей жест, падает, 2) системность организации и в то же время произвольность связи между знаком и денотатом возрастает и 3) возрастает степень общественной фиксации манеры выполнения жеста.

Тогда, согласно теории Кендона, крайне левая позиция шкалы — жестикуляция (gesticulation) обязательно требует для правильного понимания сопровождения речевым потоком, т. е. не образует системы и генерируется спонтанно. Правее в шкале располагаются жесты, которые могут заменять слова. Например, кивок вместо «да» и т. п.

На самом правом полюсе континуума или шкалы Кендона располагаются автономные жестовые системы типа языка глухонемых.

Как было показано, континуум Кендона классифицирует не типы жестов, а режимы их функционирования. Один и тот же жест может быть частью любой из систем на любом участке шкалы. Эти обстоятельства следует учитывать при классификации и исследовании любых жестовых и других семиотических систем.

Так в исследованиях А. Соломоника предлагается следующая классификация и соответственно направления исследования жестовых и других семиотических систем [4]:

1. Обыденные коммуникационные жесты: Знаки приветствия при встрече. Знаки расставания. Знаки приязни. Знаки неприязни. Знаки опасности.
2. Знаки ориентации: Название улицы, номер дома, строения, подъезда, этажа и квартиры. Образы остановок общественного транспорта и т. д.
3. Образы культурно-просветительного плана. — Музеи, театры и кинотеатры, школы и университеты, памятники и монументы, стадионы и иные спортивные сооружения, дома культуры, зоопарки.
4. Неприемлемые в данном ареале привычки и действия. — Абсцессные жесты. Нарушение порядка (в очереди, пересечение улиц в необозначенных местах, распитие и еда в садах и на улицах, бросание окурков и бумаг).

На основе вышеизложенного мы можем прийти к следующим выводам:

1. Информационные процессы, в том числе мультимедийные коммуникации, содержащие жесты и другие символы (МК) происходят в мире знаков. Это задает определенные особенности и накладывает определенные требования на исследование этих процессов, а также помогает лучше их понимать.
2. Строгие законы логики, математики, физики имеют непреходящее и неоспоримое значение. Но в знаковом мире, где реализуются МК, действуют модели, имеющие свои отличительные особенности. Эти МК реализуются в мире

универсалий, частью которых зачастую является человек, и которые допускают исключения, не влияющие на их истинность.

3. В мире МК очень эффективным и продуктивным является системный подход. Очень важно бывает понять и реализовать в форме каких универсалий и других формализмов реализуется данный ИП, который как правило является сам по себе сложной системой и является частью не менее сложной системы.
4. Для понимания, классификации и проектирования МК, происходящих в знаковом мире, а также на каких знаковых системах они реализуются следует сознательно применять законы семиотики — науки о знаках. То есть понимать и уметь формализовать их синтактику (состав знаков, их взаимосвязи —

т. е. систему), семантику — значимость и функции их частей и целого МК, а также и прагматику — доминанту и особенности их функционирования.

Литература

- [1] Рыков В.В. Обработка нечисловой информации. Управление знаниями. — М.: МФТИ, 2007.
- [2] Рыков В.В. Сервисные модели инновационной деятельности // Труды 59-й научной конференции МФТИ. — М.: МФТИ, 2016.
- [3] Рыков В.В. Чем знание отличается от информации // Независимая газета. — М.: НГ, 2017. — http://www.ng.ru/nauka/2017-09-27/11_7082_znanie.html
- [4] Соломоник А.Б. Философия знаковых систем и язык. — Минск: МЕТ, 2002.
- [5] Kendon A. *Gesture: Visible Action as Utterance*. — Cambridge, 2004.
- [6] McNeill D. *Hand and Mind: What Gestures Reveal about Thought*. — Chicago, 1992.

Когнитивные технологии в мультимедийной коммуникации — основные термины и понятия*

В. В. Рыков, Р. В. Теренин

vlad.rykov@gmail.com

Московский физико-технический институт (Государственный университет)

Коммуникация при помощи жестов и или сопровождающаяся жестами и другими знаками — так называемая мультимедийная коммуникация является быстро растущей и важной областью научного исследования. Важность эта определяется все большим числом программных проектов по реализации динамических сетевых объектов с мультипликационным сопровождением. Поэтому представляется важным знать систему понятий, а, следовательно, и терминов, их отражающих, содержащих элементарные объекты такой коммуникации, а также знание их системных взаимосвязей, и их содержания.

Ключевые слова: когнитивные технологии, универсалии, коммуникация, корпус текстов, семиотика, жестовая коммуникация, терминология, системный подход.

Cognitive Informatics in Multimedia Communication — Basic Concepts and Terminology*

V. V. Rykov, R. V. Terenin

Moscow Institute of Physics and Technology (State University)

This paper describes terminology as a crucial cognitive system and device in multimedia communication domain. Terminology helps better research, understanding and communication in this domain.

Keywords: cognitive technology, models, universals, terminology, communication, semiotics.

Любая коммуникация — в том числе и мультимедийная коммуникация (МК) есть знаковое явление. В этой области исследования есть своя система понятий и терминология, которые важно проанализировать и перечислить. Прежде всего, здесь применимы законы и терминология семиотики — науки о знаках. Семиотика делит знаки на иконические — похожие на денотат, который они означивают, а также на символы, которые реализуют чисто условную связь между знаком и денотатом. Есть также в семиотике третий вид знаков — признак. Это когда знаком является часть самого денотата. Например, дым от костра [1, 2, 4].

Поэтому следует начать с термина дискурс, как коммуникацию в самом общем виде, включающую в себя все экстрасемиотические факторы — цель коммуникации, знание культуры, в рамках которой она происходит, и т. д. Соответственно появляется элементарная дискурсивная единица (ЭДЕ), как минимальный объект МК и для представления о коммуникации или об ее

исследовании. А также — мультимедийный корпус текстов как определенным образом организованное единство, интегрированное из отдельных ЭДЕ. А также и дискурсивный маркер, описывающий ЭДЕ [3, 5, 6, 7].

Подробный словарь терминов МК есть на сайте МУРКО (www.ruscorpora.ru/instruction-murco.html). Поэтому остановимся лишь на некоторых нетривиальных терминах и понятиях. Например, при МК соответственно ее знаки (по классификации Н. И. Смирновой) могут быть — коммуникативные (замещающие речь — приветствие, прощание и т. д.), описательно-образительные знаки, теряющие смысл вне коммуникационной ситуации (знаки, отображающие форму или размер предмета и т. д.), а также модальные знаки, выражающие оценку автора предметов или людей [7].

Можно привести более экзотический термин — хиазм (буквально — перекрестие). Это синтаксическая структура, в которой одна ее часть по своей структуре симметрична другой, повторяя входящие в проти-

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20030.

воположной части ее же элементы в обратном порядке:

Например, у Лермонтова:

«Делить веселье — все готовы:

Никто не хочет грусть делить».

Невозможно в данной публикации за недостатком места описать все термины и понятия МК. Поэтому остановимся на одном из главных — когнитивный резонанс. Когнитивный резонанс — это эффект позитивной реакции на акт МК. Полученная информация соответствует когнитивной модели окружающего мира человека, активизирует адекватное решение и реализует в полной мере акт понимания [3, 5, 6, 7]. Тогда мы можем прийти к следующим выводам:

1. Терминология как система знаковых объектов, определяющих некоторую предметную область (ПО), является важной составной частью любого научного исследования или коммуникации в рамках данной ПО.

2. Следовательно, терминология это одна из форм освоения неизвестной или исследуемой ПО.
3. В науке эффективным и продуктивным является системный подход. Терминология является примером и хорошим применением системного подхода.

Литература

- [1] Рыков В.В. Обработка нечисловой информации. Управление знаниями. — М.: МФТИ, 2007.
- [2] Рыков В.В. Сервисные модели инновационной деятельности // Труды 59-й научной конференции МФТИ. — М.: МФТИ, 2016.
- [3] Кибрик А.А. Паршин П.Б. Дискурс // Энциклопедия «Кругосвет»
- [4] Крейдлин Г.Е. Невербальная семиотика. — М., 2004.
- [5] Дресвянников В.А. Жесты: обобщения и классификации // elitarium.ru
- [6] Мельников Г.П. Системология и языковые аспекты кибернетики. — М.: Сов. Радио, 1978.
- [7] Словарь терминов коммуникаций // <http://www.znaki-pr.ru/slovar.html>
- [8] Беляев И.П., Капустян В.М. Процессы и концепты. — М., 2001.

Когнитивные технологии и мультимедийные корпуса*

А. И. Климчук, В. В. Рыков, Е. В. Склярова

vlad.rykov@gmail.com

Московский физико-технический институт (Государственный университет)

В последнее время, особенно с появлением соцсетей, все большую значение начинают играть модели и универсалии, отражающие коммуникацию во всем ее объеме — не только слова, но и жесты и другие иконические знаки — так называемая мультимедийная коммуникация.

Для проведения адекватных исследований в этой области необходимы правильно построенные так называемые корпуса соответствующих текстов, построенные на базе адекватных когнитивных моделей, включающих в себя все многообразие этого вида коммуникации. Проблемы применимости этих корпусных моделей к информационным исследованиям и практике обсуждаются в данной статье.

Ключевые слова: когнитивные технологии, универсалии, коммуникация, семиотика, терминология, жестовая коммуникация, мультимедийная коммуникация, мультимедийный корпус текстов, системный подход.

Cognitive Technologies and Multimedia Corpora*

A. I. Klimchuk, V. V. Rykov, E. V. Sklyarova

Moscow Institute of Physics and Technology (State University)

The application and development of current cognitive research requires large amounts of specific resources. Mostly they are called corpora. The existing corpora are compiled either for specific problem research or to reflect the activity in general or its specific fragment. It seems that they are two distinct types of corpora or distinct objects with different design criteria. Now computer readable multimedia corpora are important instrument for various scientific studies.

Keywords: cognitive technology, models, universals, text corpus, multimedia corpus, communication, system approach, semiotics.

Речевые отношения в современном обществе все более совершенствуются и усложняются. Появляются и очень быстро становятся активными новые формы общения людей. Ярким примером этого явления может служить появление Интернет, где уже давно реализованы различные виды мультимедийных коммуникаций.

Можно сказать, что современное информационное общество приносит новый стиль в отношения людей и это требует глубокого научного изучения. Новые отношения требуют новых технологий т.к. человеку по-прежнему требуется понимать новые виды коммуникаций и уметь эффективно строить свою в новых условиях коммуникации [1, 2, 3, 5].

Каждому человеку по-прежнему требуется умение извлекать нужную ему информацию из обращенной к нему речи или прочитанного текста для принятия адекватно-

го решения в своей деловой и повседневной жизни. Также остро необходима и способность порождать уместную, эффективную устную или письменную речь — как на новых носителях, так и в традиционных условиях деловой и бытовой коммуникации.

Каждая коммуникация должна находиться в рамках культурных традиций данного общества. Трудно себе представить, чтобы деловое письмо или научная статья, написанные вне рамок соответствующих культурных традиций, могли бы быть достаточно эффективными и даже просто достичь адресата.

Можно также сказать, что по-прежнему остается актуальной проблема оптимизации обучения построению эффективной коммуникации. Существует мнение, что эта проблема сейчас стоит даже как никогда остро. Есть много различных подходов к решению этой проблемы. Некоторые подходы решают эту проблему в рамках существующих традиций, развивая их в соответствии с изменившимися обстоятельствами.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20030.

Одна из таких традиций основывается на обучении и усвоении образцовых для данной культуры и условий коммуникации объектов. По сути человеку, желающему научиться эффективно строить свою коммуникацию в определенных условиях, предлагается корпус образцовых или учебных мультимедийных текстов, на базе которых он проходит соответствующее обучение. Другими словами, обучающемуся предлагается корпус образцовых текстов, ориентированных на усвоение и использование их как прототипа для последующих речевых действий в определенных условиях коммуникации.

Для проведения адекватных исследований в этой области необходимы правильно построенные так называемые корпуса соответствующих текстов, построенные на базе адекватных когнитивных моделей, включающих в себя все многообразие этого вида коммуникации. Проблемы применимости этих корпусных моделей к мультимодальным коммуникациям — их теории и практической реализации обсуждаются в данной статье.

Корпус текстов — это определенным образом организованное множество, элементами которого являются тексты. Организация корпуса может быть самая разная — в зависимости от прагматических целей его создателя или пользователя. Тексты, которые суть составляющие элементы корпуса, могут представлять собой целое оригинальное словесное произведение или какую-либо его часть. Однако, элементами корпуса могут быть не только тексты. Существует понятие дискурс, включающее в себя все нетекстовые объекты, участвующие в коммуникации — жесты, психологические, культурные, социальные и психологические объекты (filologia.su/diskurs/). Корпус, включающий в себя такие объекты это уже мультимедийный корпус (МК).

Тогда элементами такого МК будут более сложные объекты — так называемые элементарные дискурсивные единицы (ЭДЕ). Они включают в себя все виды вышеупо-

мянутые объекты как свои элементы. Получаются так называемые мультимедийные корпуса текстов. Таких корпусов уже очень много. Поэтому мы здесь рассмотрим самые массовые и авторитетные из них, которые чаще всего используются в научных и практических применениях [4, 6].

Как уже говорилось, подробные описания этих корпусов есть на указанных ниже сайтах. Поэтому остановимся на описании основополагающих принципах их создания и применения.

За недостатком места остановимся только на их кратком описании — т. к. более полная информация об этих корпусах есть на упоминаемых ниже их сайтах (адресах).

Прежде всего — это Русский мультимедийный дискурс, разработанный в Институте языкознания РАН (multidiscourse.ru). Создатели этого корпуса включили в его ЭДЕ записи не только слов, но и интонации, мимику, движения глаз и т. д. При помощи этого корпуса проблемная группа ставит целью поддержание различных научных и практических применений.

Другим авторитетным корпусом является так называемый МУРКО — Мультимедийный русский корпус (ruscorp.org.ru). ЭДЕ в МУРКО является так называемый кликст — клип плюс текст. Сейчас в МУРКО более 5 млн кликстов.

Можно также упомянуть корпус русского жестового языка (rsl.nstu.ru). А также Русский эмоциональный корпус (REC), который используется в робототехнике. А также так называемый LiveCorpus, на базе которого можно научиться аннотировать и вообще создавать другие корпуса.

После рассмотрения проблем построения мультимедийных корпусов текстов (МК) можно прийти к следующим выводам:

1. Проблема репрезентативности любого корпуса проблемной информации является весьма актуальной для любого исследования. Это существенно важно как для теоретических изысканий, так и для прикладных задач.

2. Репрезентативность как доказанное качество, присущее сконструированному корпусу и, конечно, МК определяет достоверность полученных на нем результатов. Проблему можно рассматривать как проблему адекватного отражения, адаптации или интеграции больших массивов информации в существенно меньший по объему корпус данных.
3. Для решения этой проблемы в области исследования коммуникации предлагается модель МК.
4. Для построения МК реализуется модель (парадигма) объектно-ориентированного программирования (ООП). Здесь реализуются такие классические методы ООП как наследование, полиморфизм, инкапсуляция и другие.
5. Анализ этой проблемы показал, что деятельность по составлению или компиляции МК чрезвычайно разнообразна и, следовательно, довольно неоднородна. Во-первых, это определяется разнообразием задач, стоящих перед составителем корпуса. Во-вторых, мы наблюдаем здесь обычную полисемию термина — в данном случае термина «корпус текстов». Или, другими словами, корпус текстов — родовое понятие множества довольно разнообразных объектов. Зачастую эти объекты связывает только общее родовое имя.
6. В корпусной лингвистике принято определение корпуса, основывающееся на четырех признаках [6]. Этими основными содержательными признаками, которые определяют специфику этого своеобразного рода словесного единства являются следующие: 1) расположение корпуса на машинном носителе, 2) стандартизованное представление мультимедийного материала на этом машинном носителе, позволяющем применять стандартные программы его обработки, 3) конечный размер, 4) репрезентативность как результат особой процедуры отбора.

Литература

- [1] Рыков В.В. Обработка нечисловой информации. Управление знаниями. — М.: МФТИ, 2007.
- [2] Рыков В.В. Сервисные модели инновационной деятельности // Труды 59-й научной конференции МФТИ. — М.: МФТИ, 2016.
- [3] Рыков В.В. Корпус текстов как отражение состояния русского языка // Труды Международного конгресса «Русский язык: исторические судьбы и современность». — Москва: МГУ, 2001.
- [4] Кибрик А.А. Паршин П.Б. Дискурс // Энциклопедия «Кругосвет».
- [5] Крейдлин Г.Е. Невербальная семиотика. — М., 2004.
- [6] Дресвянников В.А. Жесты: обобщения и классификации // elitarium.ru
- [7] Мельников Г.П. Системология и языковые аспекты кибернетики. — М.: Сов. Радио, 1978.

Когнитивные модели в инновационной деятельности*

Е. С. Максимов, В. В. Рыков

vlad.rykov@gmail.com

Московский физико-технический институт (Государственный университет)

Рассматриваются различные способы организации и оптимизации инновационной и других видов деятельности на основе когнитивных моделей. В процессе своей жизнедеятельности люди взаимодействуют с различными объектами — как с объектами реального окружающего мира (материальными объектами), так и с идеальными объектами, а также и с объектами знаковой природы. То есть с тремя видами объектов. Люди придумали много приемов, технологий, теорий и инструментов как оптимизировать эту деятельность. С середины XX века начала бурно развиваться информатика, которая резко повысила эффективность работы со знаковыми объектами. В последние годы все более часто стали применяться когнитивные модели и технологии, которые имеют дело с так называемыми когнитивными объектами или обрабатывают знание. Применение этих моделей рассматривается в этой статье.

Ключевые слова: когнитивные технологии, модели, универсалии, коммуникация, семиотика, инновация, сервис.

Cognitive Models in Innovation Activity*

E. S. Maksimov, V. V. Rykov

Moscow Institute of Physics and Technology (State University)

This paper describes the crucial approaches in cognitive technologies. The info processing runs in the world of signs. It means that they are based on semiotic universals. They are often based on communication. So these universals are available to them. This paper discusses the problems of cognitive models application to info processing and presumably to innovation activities. The important feature of proposed technologies is its availability for an engineer.

Keywords: cognitive technology, models, optimization, innovation activity, service, universals, communication, semiotics.

По аналогии с тем, что материальный мир можно разложить на молекулы, а затем и атомы, то и человеческую деятельность можно разложить на элементарные процессы — назовем их бизнес процессами (БП). Но и сам БП, по аналогии с молекулой, можно рассматривать как совокупность трех видов объектов — входные объекты, выходные объекты (результат БП) и сам процесс преобразования входных объектов в выходные [1, 5, 6].

Оказывается, что БП можно классифицировать по элементарным когнитивным признакам — что мы знаем или не знаем о его составных частях — входе, выходе и о самом процессе. Существует классификация, основанная на этих когнитивных признаках. А именно — если знаем о входе БП — то это единица (1), а если не знаем, то ноль и т. д. И тогда элементарный БП кодируется тройкой битов — 0 или 1. Тогда,

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20030.

например, 000 означает, что мы не знаем о начале БП, что происходило и чем закончилось. Можно самому попробовать рассмотреть разные комбинации, однако приведем несколько примеров.

Так, например, код 011 означает, что нужно определить с чего начался определенный БП, который привел к известному результату. Или — 101 — изобрести такой БП, который из заданных начальных объектов привел бы к определенному результату. Более того, методами инфографики можно изобразить входные и выходные объекты БП в виде прямоугольников. Из входных объектов стрелки будут направлены к овалу — самому процессу. А справа из этого овала стрелки будут указывать на выходные объекты (тоже прямоугольники). Практика показывает, что такое изображение методами инфографики деятельности даже небольшого рабочего коллектива приводит к положительным результатам [1, 2, 3, 5, 6].

Тогда далее проблема сводится к методам или моделям интеграции таких элементарных БП в более сложную систему. Одной из таких моделей является так называемая сервисная модель. В процессе инновационной деятельности все чаще в научно-производственных коллективах и конкретно также и в МФТИ вместе оказываются люди совершенно разных специальностей, а зачастую и культур. Одной из проблем стала оптимизация их взаимодействия, результатом решения которой стала разработка термина и стандарта под названием сервис, ничего не имеющего общего с бытовым понятием сервиса. Понятие и стандарт сервис были описаны в Белой книге, написанной в процессе обсуждения в Кембридже при участии различных специалистов [7].

Интересно, что за работы по описанию и внедрению понятия сервис несколько ученых получили Нобелевскую премию по экономике в 2009 году. Одним из базовых понятий теории термина является теория взаимодействия, восходящая к понятию регламент. Оказалось, что в некоторых научно-производственных коллективах при работе даже нежелательно общение между некоторыми членами коллектива. Конечно, и без этого эффекта видна культура регламента. Действительно, невозможна передача знаний даже школьникам, если они шумят во время урока. Аналогичный эффект можно наблюдать и для научных семинаров, лекций и конференций.

Однако современность потребовала более детальной проработки этой традиции и, как указывалось выше, была создана наука — теория сервиса (SSME — service science management engineering). Курсы теории сервиса читаются во многих вузах мира. Нетрудно видеть, что элементами сервисной когнитивной системы являются прежде всего люди.

Но тогда, если рассматривать множество объектов, подлежащих научному исследованию в рамках того или иного инновационного проекта, то необходимо обозначить такую триаду научной деятельности:

1. Методика — статистические или экспериментальные приемы выявления в исследуемых объектах релевантных типологических признаков или характеристик.
2. Метод — положенные в основу когнитивные процессы выведения новых знаний о типологических особенностях изучаемых объектов.
3. Теория или концепция, явно или неявно представленная в виде классификационной схемы. Это проявляется через систему понятий, особенностей признаков каждого из таксонов и таксономических соотношений между классами.

То есть нужна такая система интеграции людей, их знаний и объектов исследования, которая не зависела бы от алгоритмов ее создания. Однако представленная выше сервисная методология или концепция вполне адекватно может быть приложима к особому типу объектов — совокупности людей, участвующих в инновационном проекте (ИП), их взаимоотношений, объектов и результатов их научных исследований. Сервисная система определяется как динамическая конфигурация различных ресурсов (людей, технологий, организации и информации), формализующая и оптимизирующая их взаимодействие и результаты этой деятельности. Не следует думать, что это невозможно в рамках описанной выше триады. Это уже триада более высокого уровня, где объектом исследования является уже исследовательская деятельность.

Сервисная система не противопоставляет себя процессам самоорганизации в исследовательских группах — а только вносит необходимые алгоритмы и критерии оптимизации [2].

Таким образом, сервисная система позволяет выполнить и оптимизировать следующие процессы [1, 2, 3, 7]:

1. Интегрировать в данный инновационный проект (ИП) все необходимые объекты — участников ИП, их коммуника-

- ции, их знания, объекты исследования этого ИП.
2. Синхронизовать и оптимизировать коммуникацию участников и их действия над объектами ИП.
 3. Решить проблему именованности (vocabulary integration). Здесь имеется в виду часто встречающееся явление, заключающееся в том, что разные специалисты могут по-разному называть одни и те же объекты.
 4. Эти преимущества дают возможность руководителям ИП смелее развивать свои различные междисциплинарные исследования и бизнес проекты.
 5. Сервисная система (SSME), таким образом, представляет собой оптимальную и типовую технологию интеграции некоторого фрагмента человеческой деятельности в единую систему.
 6. Эти преимущества позволяют осуществлять как локальную оптимизацию существующих ИП, так и внести радикальные изменения в стратегию ИП.

Литература

- [1] Рыков В.В. Обработка нечисловой информации. Управление знаниями. — М.: МФТИ, 2007.
- [2] Рыков В.В. Сервисные модели инновационной деятельности // Труды 59-й научной конференции МФТИ. — М.: МФТИ, 2016.
- [3] Рыков В.В. Информация, классификация и системный подход // Труды 57-й научной конференции МФТИ. — М.: МФТИ, 2014.
- [4] Альтшулер Г.С. Алгоритм изобретения. — М., 1969.
- [5] Беляев И.П., Капустян В.М. Процессы и концепты. — М.: СИМС, 1997.
- [6] Иванов А.И., Красюк Т.В. Синергетические модели самоорганизации в малых группах. — М.: Спутник, 2010.
- [7] Succeeding through service information. White Paper. — Cambridge: UCIM, 2008.

О проблемах имиджевого возвышения России*

Ю. Н. Голубчиков¹, С. В. Клименко²
golubchikov@list.ru

Москва, ¹Географический факультет МГУ имени М. В. Ломоносова
Протвино, ²Институт физико-технической информации

Рассматривается положение России в русскоязычном, англоязычном и китайскоязычном секторах Интернета. Выражается обеспокоенность недостаточностью имиджевого продвижения России в мировом информационном, культурном и туристическом пространстве. Обращается внимание на глобальную роль биосферно-терапевтического значения России. Из всех отраслей экономики больше всех заинтересована в возвышении своей страны ориентированная на внутренний въезд туристическая индустрия.

Ключевые слова: экология, туризм, Россия, имидж, пространство.

On the problems of image promotion of Russia*

Yu. N. Golubchikov¹, S. V. Klimenko²

¹Faculty of Geography, Lomonosov Moscow State University
Protvino, ²Institute of Physics and Technology

The situation of Russia in the Russian-speaking, English-speaking and Chinese-speaking sectors of the Internet is considered. Concern is expressed about the lack of image promotion of Russia in the global information, cultural and tourist space. Attention is drawn to the global role of the biosphere-therapeutic significance of Russia. Of all the branches of the economy, the tourism industry oriented in inner tourism is most interested in prominence of the country.

Keywords: ecology, tourism, Russia, image, space.

Ныне в мире вновь возрождается геополитическое мышление, связывающее возможности государств с расширением и перераспределением пространства. Значение его взрывообразно растет вместе с демографическим взрывом. Страны бедные, но с интенсивным демографическим ростом, надеются на заселение экономически преуспевающих стран. У тех, в свою очередь, поиски экологически чистых средств становятся основным занятием экономистов, политиков и ученых. К ним присоединяется сильный и демографией и экономикой Китая. Все большее внимание тех и других и третьих привлекает огромность и незаселенность России.

Мировой общественности рисуется холодная и загрязненная страна, место частых техногенных и природных катастроф, хотя и располагающая обширными запасами нефти и газа. Постоянно указывается на нерадивость и бесхозяйственность в использовании российских пространств. Ставится

вопрос об ограничении прав государств, не способных распорядиться выпавшими на их долю природными богатствами «нашего дома — космического корабля Земля».

В то же время в России расхожими стали представления, что в наш век информационных технологий ценность пространства и природных ресурсов неуклонно снижается. «Пространство, — твердят нам — многого не стоит, а стоит что-то стоящее, например, компьютер или новая марка автомобиля». Складывается представление о больших пространствах как источнике всяких бед. Оттого, наверное, кое-какие россияне не усматривают в разрушении страны особой беды. «Живут же маленькие Нидерланды!», — восклицают они. «Мы отдаем нашим необъятным просторам больше, чем получаем» — вторят им некоторые российские ученые и политики.

Согласно данным многочисленных социологических опросов о привлекательности России, страна неизменно оказывается в середине или конце списка. Студенты из 17 стран поставили Россию по привлека-

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 16-07-00955, 18-07-20030.

тельности на 166-е место среди государств и территорий мира [9]. По другим рейтингам Россия занимает 59 место из 133 стран, тогда как не имеющая никакой истории Канада стоит на 5 месте [10]. В туристических поездках россияне также отдают предпочтение посещению зарубежных стран. Это отчетливо прослеживается в пространственном распределении в Твиттере 4,8 млн. ссылок на русскоязычные комментарии (в т. ч. к фотографиям), взятых с 21 мая по 9 августа 2017 года (рис. 1). На русскоязычный сегмент пришлось лишь 1,2% от всех ссылок. Они косвенно отражают степень туристической активности.

Можно спорить о достоверности изучения распределения туристов по Твиттеру. Только не следует упускать из виду недостоверность туристической статистики. Так, согласно официальному Докладу о состоянии и развитии туризма в Российской Федерации в 2015 году [5] численность туристических посещений Ямало-Ненецкого округа составила 484 тыс. чел. Магаданскую область за тот же год посетило всего 1 тыс. чел. Но по численности геотегов эти субъекты оказались сопоставимы. Известно, что более половины арктических туристов в РФ приходится на Мурманскую область (около 350 тысяч) и она действительно отражена в соответствии с этим числом на Твиттере и в Докладе [5].

Еще меньше отражена Россия на англоязычных ресурсах Твиттера (рис. 2), на которые приходится половина всех ссылок (165,2 млн из 387,4 млн. всех ссылок).

Все англоязычные ссылки сконцентрированы в основном вокруг Москвы и Санкт-Петербурга. Восточнее Петербурга они образуют резкий спад. Даже китайские туристы, вопреки бытующему в СМИ мнению, очень слабо отражены в Твиттере (рис. 3).

В деятельности каждого производства огромное значение придается рекламе продукции. А у самой огромной страны ни одно ведомство не озабочено ни её рекламированием, ни украшением, ни имиджевым

продвижением. Из всех отраслей экономики больше всех заинтересована в возвышении своей страны ориентированная на внутренний въезд туристическая индустрия. Очевидно, что благодаря лишь туризму Россия стала одним из крупных финансовых доноров для многих стран мира. На цели рекламирования России в мировом туристическом пространстве выделяется 5 млн. евро в год, а по европейским странам в среднем около 30 млн. евро (в Турции — 112 евро) [4]. На этот заказ ежегодно выделяются средства, входящие в число приоритетных расходных статей государственного бюджета (таблица 1). Для авторитета государства конструирование притягательного портрета сегодня становится поважнее укрепления военно-экономического могущества. Хотя многое сделано по повышению привлекательности России после того как Попечительский совет РГО возглавил В. В. Путин, а президентом РГО стал С. К. Шойгу, по-прежнему очень велики расходы на экологическую антирекламу страны.

Лучшее что можно сделать для конструирования России как великой туристической державы это прививать любовь к ней. А любовь начинается с восхищения. У нас есть чем восхищаться. Это наши имперские атрибуты, которых мы как бы стесняемся, наше положение на перекрестке цивилизаций, великий русский язык, уникальная русская культура.

В этом плане следует заявить о русском народе как о крупнейшем народе мира. По степени распространённости русский язык занимает четвёртое место в мире после английского, китайского и испанского языков. Но далеко не все, считающие английский родным, идентифицируют себя как англичане. Американцы США не образуют единого антропологического типа и делятся на несколько расовых групп. Американский народ — понятие такое же искусственное, как и советский народ. Испаноязычные латиноамериканцы тоже не отождествляют себя с испанцами. Китайцы вообще не говорят на одном языке. В Китае жи-

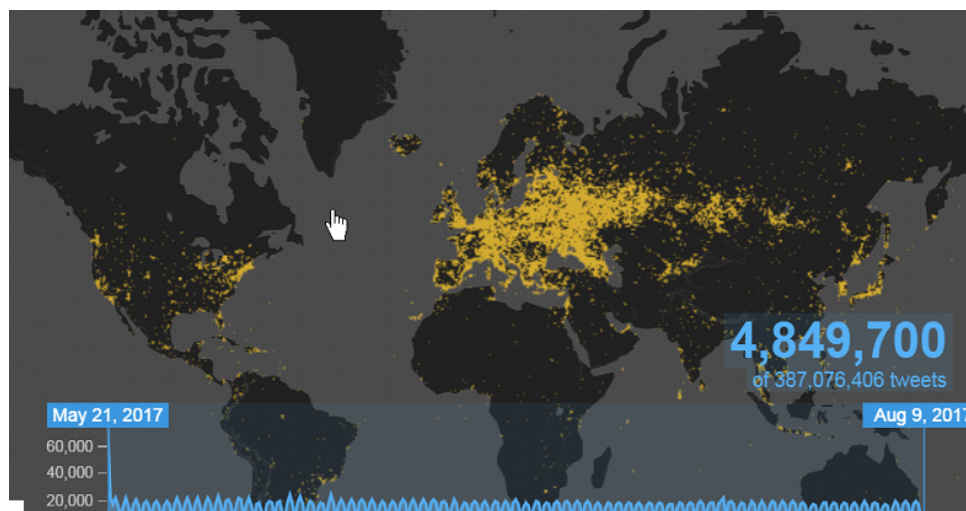


Рис. 1: Россия на русском языке

Пространственное распределение в Твиттере 4,8 млн. ссылок на сообщения и фотографии в интернет-сообществах, блогах и на различных сайтах (твиты и посты), которые система интерпретировала, как русскоязычные с 21 мая по 9 августа 2017 года. Отчетливо проявляется концентрация ссылок к Средней и Южной Европе, Южной Сибири, Приморью
<https://www.mapd.com/demos/tweetmap/>

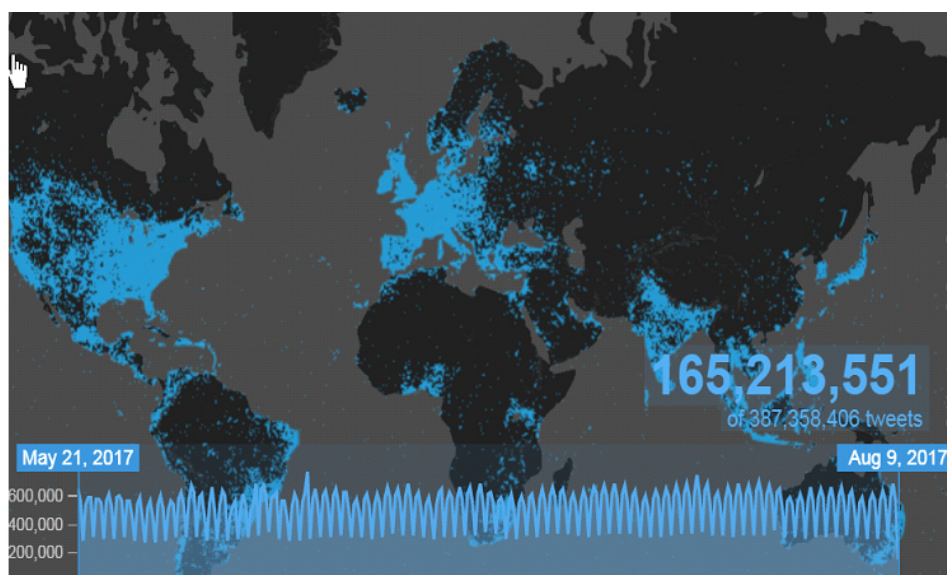


Рис. 2: Россия на английском языке

Пространственное распределение в Твиттере 165,2 млн. ссылок на сообщения и фотографии в интернет-сообществах, блогах и на различных сайтах (твиты и посты), которые система интерпретировала, как англоязычные с 21 мая по 9 августа 2017 года. Север России лежит практически во мгле
<https://www.mapd.com/demos/tweetmap/>

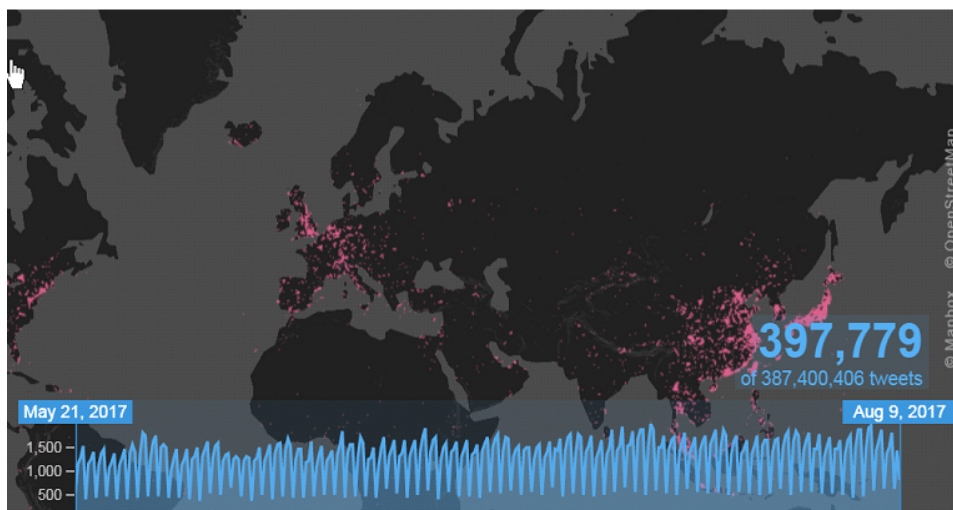


Рис. 3: Россия на китайском языке

Пространственное распределение в Твиттере 398 млн. ссылок на сообщения и фотографии в интернет-сообществах, блогах и на различных сайтах (твиты и посты), которые система интерпретировала, как китайоязычные (в т. ч. к фотографиям) с 21 мая по 9 августа 2017 года

<https://www.mapd.com/demos/tweetmap/>

вет 50 народов и северные китайцы совершенно не понимают южных. Связывает их воедино иероглифическая письменность. Ее знаки записывают не звук, а смысл. Ещё менее консолидированы индусы (16 государственных языков Индии), индонезийцы (726 «живых» языков) и нигерийцы (505 языков) [14]. А вот русские не только говорят на одном и том же русском языке от Львова до Владивостока и от Ташкента до Норильска, но и обладают единой культурой, антропологией и психологией.

Россия — геополитический продолжатель самых крупных в истории межконтинентальных империй: держав гуннов, Чингисхана, царей и Советского Союза. Для Индии Россия остается таинственной прародиной Гипербореи (Арктиды). Конечно, это антинаучно. Но о чудовище в Лох-Нессе тоже неизвестно, есть оно там или нет, но миллиарды долларов крутятся. Хотя с помощью антропологического материала и ДНК-генеалогий выявлено, что Север Русской равнины — это единственный регион мира, где живет население, не испытывавшее с эпохи верхнего палеолита мощных генетических и культурных влияний [6]. Отсю-

да уходили, но никто не мигрировал сюда массово.

Пространства России служили, по образному выражению Гоголя, вулканом народов, заселявших Европу. Вся Германия и вся Скандинавия были заселены нашими предками о чем говорит их топонимика. Вышедшие из степей Причерноморья готы основали королевства Испании. Магьяры и болгары до сих пор хранят воспоминания о своей исторической прародине в Приуралье.

Через Византию Россия наследует античность Древней Греции. Сама формула «Москва — третий Рим» напоминает о верности России первоначальным канонам христианства. Она могла бы стать в этом отношении «Меккой» православия, к которому в мире пробуждается все больший интерес. Атеистическая и материалистическая Европа ищет отсюда возвращения к истокам. Все большее значение обретает православие для воскресшего от культурной революции Китая. Китайцев также интересует Россия как колыбель Октябрьской революции, страна связанная с именами Ленина, Сталина, Мао Цзе Дуна. А китайцы — это наш главный туристический ресурс.

Россию постоянно обновляла бурная история катастроф. Хотя прошлое нас скорее разделяет. История России систематически искажалась и сводилась от 1000-летней к 70-летней, то к 300-летней или непонятно какой. В дореволюционной истории идеологией являлось православие, в советской — коммунизм. Только у нас каждое поколение сносит памятники предшествующего. В настоящем также немало разделяющего. По политическим убеждениям мы можем быть правыми, левыми, коммунистами, анархистами, демократами, либералами. Но связывает всех нас география России. Если какая-то ее часть будет отринута, то обратно ее уже не вернут. В результате наиболее обеспечивают политическую преемственность страны именно географические атрибуты. Как ни у одной из стран мира, география может служить философией России и ее национальной идеей не только из-за обширности пространств, но еще из-за размытости исторических пределов. «Сегодня нам как никогда нужна национальная география! География, которая способствует реализации национальных интересов», — отмечает Президент РГО С. К. Шойгу [18].

Наши предки оставили нам великое достояние. Во всей всемирной истории невозможно указать другой подобный пример окончательного завоевания столь обширных пространств с таким сложным этническим составом, в столь короткий срок и таким ничтожным количеством людей. Русские казаки прошли всю Сибирь от Урала до Тихого океана за 60 лет, тогда как англо-европейцам для того, чтобы освоить территорию Северной Америки от Атлантического до Тихого океана, понадобилось около 350 лет. Еще через 100 лет русские утверждаются на Аляске. На границе Аляски и Канады русское племя встречается с англосаксонским, которому удалось к середине XIX в. пройти весь американский материк. Две ветви европейцев, направившись в разные стороны почти из одного центра, вновь сошлись на другом конце земного шара. История их народов — есть история коло-

низаций. Но отношение к этой истории разное.

Фронтирность стала национальным триумфом американской истории. Создание же Империи русскими землепроходцами не стало базовым компонентом отечественной истории. Нет у нас какой-либо «землепроходческой» атрибутики, подобной ковбойской или золотоискательской [3]. Нет атрибутики «Дикого поля» как степной границы между Востоком и Западом, Севером и Югом [2]. И подавно ничего не найти о западном фронтире, утвердившемся в Черновцах, Львове, Риге после Великой Отечественной войны.

Россия охватывает и соединяет Восток и Запад, Север и Юг, соприкасается с главными мировыми цивилизациями. Такому положению страны отвечает ее государственный герб. Многие полагают, что Россию неправомерно соотносить ни с Европой и ни с Азией. Это особый географический и исторический мир, соразмерный и равноправный как Европе, так и Азии.

Вопреки широко внушаемым в определенном секторе научной литературы и публицистики представлениям об уничтоженной природе России, ни одна из стран не сохранила за собой столь ненарушенных пространств, как Россия в своих северных и восточных владениях. Почти половина ее пространств неотягощена антропогенной и техногенной нагрузкой. Это значит, что, пролетая над ними на самолете или разглядывая снимки из космоса, не удастся обнаружить никаких значимых следов человеческой деятельности — населенных пунктов, дорог, водохранилищ, линий электропередач, возделанных полей, даже огоньков среди ночи не видно. В Европе таких ненарушенных площадей географ Н. Н. Клюев [7, 8] насчитывает всего 2,8%.

Ни одна из стран не обладает такой огромной площадью особо охраняемых природных территорий, как Россия. Государственные заповедники, национальные и природные парки, заказники, памятники природы

занимают 12% от общей площади страны, столько же, сколько и пашня [12].

В России сосредоточено 25% всех мировых запасов леса, в 3 раза больше, чем в США и Канаде, вместе взятых, Лес занимает 69% территории страны. Леса России образуют крупнейший в мире сравнительно единый массив естественных экосистем. Они заметно улучшают состояние российских городов. Ни одна из столиц мира не имеет в своем ближайшем окружении столь хорошо сохранившихся лесов, как Москва. Настоящие леса прямо заходят а ее черту в виде Лосинога острова.

Именно российские холодные консервирующие пространства являются главным поставщиком кислорода на планете. Тропические же леса больше поглощают кислород на процессы разложения и гниения, чем его производят. Теоретически страны должны платить не за свой углекислый газ, от которого никто еще не умер, а получать доходы от других стран за поставку кислорода. К тому же российский кислород особого качества, насыщенный отрицательными ионами. А это самое важное и доступное средство исцеления человека географической средой. Самое замечательное и бесплатное лекарство.

Очень насыщен отрицательными ионами воздух после грозы, когда он напоен тысячами ароматов цветов, листьев, травы. К ионизирующим факторам относятся также: так называемые тихие электрические разряды у крон высоких деревьев и на вершинах гор, возникающие при больших значениях напряженности электрического поля атмосферы; распыление и разбрызгивание воды у горных рек и водопадов, фонтанов, во время прибоев у побережья морей и океанов. Очень ионизирован отрицательными ионами и полезен кислород в хвойном лесу [16, 17].

России принадлежит половина хвойных лесов планеты — мощнейших ионизаторов. До 40% площади страны занимают горы, с их чистейшими пузырящимися от из-

бытка кислорода водотоками. Существенно обогащает воздух отрицательными ионами кислорода и самая протяжённая береговая линия России с очень сильными приboями и приливами на изрезанных побережьях. Наконец, в стране немало дней с грозами. На фоне гигантского количества публикаций-диссертаций по необходимости снижения концентрации углекислого газа в связи с предполагаемым потеплением климата, имеется ничтожно мало научных статей по поставке Россией атмосферного кислорода, а по отрицательным его ионам, похоже после А. Л. Чижевского вообще никто ничего ни считал и ни писал.

После кислорода, самым важным веществом для человека является вода. Наиболее целебны талые снеговые и ледниковые воды. И в этом отношении ни одна территория не сопоставима с Севером России по объемам и продолжительности воздействия талых снеговых и ледниковых вод [1].

Наконец, целебна пища самых экологически чистых в мире российских сред. Практически все отечественные продукты питания отличаются высоким качеством. Внесение минеральных удобрений на 1 га посево составляет в мире 100 кг/га (в Великобритании — 234, в Китае — 648), а в РФ — 42 кг/га [8]. Если бы не настойчивые утверждения «зеленых» о загрязнении, в том числе радиоактивном, отечественное продовольствие стало бы вполне конкурентоспособным на мировом рынке.

С позиции ВВП завтрашнего дня сохранившаяся природная среда и красота территории значат не меньше богатств ее природных ресурсов. Как писал французский эколог Ф. Сен-Марк [17], сырье можно заменить другим сырьем, ненарушенную территорию заменить нечем. Ненарушенные территории в этом отношении бесценны для географически близких к России густонаселенных государств Западной Европы, Азии и Америки. А там начинает осознаваться, что высокая плотность и скученность населения — главный фактор любых заражений. Странно выглядит ситуа-

ция, когда пребывая год в скученном мегаполисе, человек устремляется на отдых в не менее переполненное людьми курортное побережье или в столь же многолюдные туристические центры. Поправлять здоровье следовало бы как раз в малолюдной или даже безлюдной экологически чистой сред. Постоянно растет спрос на путешествия, ландшафтный туризм, ландшафто-терапию, природное целительство.

Отмечается затухание стандартных пакетных туров, пляжного и шопингового туризма. Расширяются виды индивидуального туризма. Люди путешествуют за обретением нового опыта и смысла, физического укрепления и овладения техниками выживания. За последние десятилетия, благодаря расширению парка малой авиации, спектр видов спорта существенно расширился. Появились такие мобильные его виды как фрирайдинг, скайдайвинг, фриклаймбинг, парапланеризм, рафтинг, виндсерфинг и пр. Объединяет новые виды спорта вовлеченность человека в стихии дикой природы. Где как не в России найти для них такой простор. В этом отношении практически все наши приарктические пространства представляют собой уникальный психологический санаторий и наилучшее убежище от технологической цивилизации. Места с такими богатствами могли бы стать нашим национальным достоянием, но у нас они окружаются депрессивными поселками и обезлюдившими селами.

России принадлежит 15% всех практически не затронутых деятельностью человека территорий земной суши, 20% мировых запасов пресных вод. У нас сохранилось четверть мест внеледниковой земной суши, где еще не ступала нога человека. И все эти богатства сосредоточены в основном на Российском Севере. Глобальная их роль в том и состоит, что они своего рода биосферно-терапевтический резерват земного шара. Нам бы следовало бы капитализировать этот наш уникальный ресурс. Однако сегодня все русское и российское стало чем-то немодным.

Около 2-3 миллиардов населения мира с сотнями миллионами потенциальных туристов сосредоточены совсем под боком России. Прежде всего это быстрорастущий Китай с совершенно сведенными лесами и экологически неблагоприятными условиями жизни даже в сельской местности. Отсюда в недорогие виды туризма легко вовлечь широкие слои среднего класса населения, неприспособленного к санитарно-гигиеническим условиям и тянущегося к уединенному отдыху в экологически чистой природе и исторически близких, но малоизвестных пространствах при недорогом переезде. В Китае действует программа государственной поддержки туризма с компенсацией до 500 долларов затрат на человека.

Как пишет замечательный сибирский писатель А. К. Омельчук [11]:

«Надо сформулировать:

Россия — главная страна на этой планете,
Если даже это не имеет никакого отношения к истине.

Россия — главная страна на Земле».

Литература

- [1] Алексеев В.Р. Талая вода — криогенный ресурс планеты // География и природные ресурсы. 2012. № 1. С. 14–24.
- [2] Афанасьев М.В. Водные путешествия по плато Путорана. М.: Компания Спутник +, 2007. 248 с.
- [3] Белаш Н.Ю. Фронтьер в Америке. Фронтьер в России? // География, 1996, №5, С.12–13.
- [4] Быстров С.А. Региональные аспекты налоговой политики: адаптация мировой туристской практики к регионам Российской Федерации // Вестник СПбГУ. Сер. 7. 2016. Вып. 1. С. 90–102.
- [5] Доклад о состоянии и развитии туризма в Российской Федерации в 2015 году. М.: Министерство культуры Российской Федерации, 2016, 212 с.
- [6] Клесов А.А. Происхождение славян. ДНК-генеалогия против «норманнской теории». М.: Алгоритм, 2013, 512 с.
- [7] Клюев Н.Н. Россия на экологической карте мира // География, 2001, №47, С. 3–6.
- [8] Клюев Н.Н. Экологическая география России. М.: Русское слово, 2017, 96 с.
- [9] Колосов В.А., Зотова М.В. Геополитическое видение мира российскими гражданами: почему Россия не Европа? // Полис, 2012, №5, С. 170–186.

- [10] Матвейчев О.А. Что делать России? Прорывные стратегии третьего тысячелетия. М.: Эксмо, 2012, 352 с.
- [11] Омельчук А.К. Високосный год. 366. Тюмень: Инфо-плюс, 2013, 366 с.
- [12] Родзевич Н.Н. Экологическая самодостаточность России // География в школе, 2002, №1, С. 3–9.
- [13] Сен-Марк Ф. Социализация природы. М.: Прогресс, 1977. — 435 с.
- [14] Смирнягин Л.В. О региональной идентичности // Пространство и время в мировой политике и международных отношениях. Материалы 4-го Конвента РАМИ. В 10 томах. Т. 2. Идентичность и суверенитет: новые подходы к осмыслению понятий. Рос. ассоциация междунар. исследований. М.: МГИМО-Университет, 2007. С. 81–107.
- [15] Тен В.В. Народы и расы. Происхождение. СПб.: Инсайт, 2013, 448 с.
- [16] Чижевский А.Л. Атмосферное электричество и жизнь // Земля во Вселенной. М.: Мысль, 1964. С. 422–442.
- [17] Чижевский А.Л. Аэроионификация в народном хозяйстве. 2-е изд., сокр. М.: Стройиздат, 1989. 488 с.
- [18] Шойгу С.К. Речь Президента РГО С.К. Шойгу на Арктическом форуме 21 сентября 2010 г. // <http://www.rgo.ru/2010/09/shojgu-razvitie-nacionalnoj-geografii-dolzno-sluzhit-interesam-rf/>

Разработка и реализация программного обеспечения «Геоконвертор» в составе геоинформационной системы «Геоанализ»*

М. И. Судейкин, А. А. Афанасьев
MSudeykin@ermasoft.ru

Московский физико-технический институт (государственный университет)

В данный момент в составе программного комплекса «Геоанализ» подготовлено большое количество источников данных (географических карт), предназначенных для работы в более ранних версиях ПФК. Результатом же данной работы является разработка программного обеспечения ПО «Геоконвертор», которое позволит конвертировать старые источники данных для работы с новой версией ПФК. Важной составной частью ПО являются алгоритмы предварительной обработки геоданных, (различные алгоритмы генерализации, алгоритмы формирования названий и подписей геообъектов и т. д.).

Ключевые слова: программное обеспечение, геоинформационная система.

Development and implementation of the software “Geoconvertor” in the geoinformation system “Geoanalysis”*

M. I. Sudeykin, A. A. Afanasiev

Moscow Institute of Physics and Technology (State University)

At the moment, a large number of data sources (geographic maps) designed for work in earlier versions of PFC have been prepared as part of the Geoanalysis software package. The result of this work is the development of software “Geoconvertor”, which will allow you to convert old data sources to work with the new version of PFC. An important part of the software are algorithms for preliminary processing of geodata, (various algorithms for generalization, algorithms for the formation of names and signatures of geoobjects, etc.).

Keywords: software, geoinformation system.

Введение

В настоящее время, в связи с постоянно растущим объемом картографических данных, а также в связи с ужесточением требований пользователей к быстродействию системы, возникла необходимость разработки новой версии ПФК «Геоанализ», которая могла бы работать значительно быстрее предыдущей.

Основным преимуществом новой версии ПФК «Геоанализ» по сравнению с предыдущими версиями является качественно новая схема хранения геоданных. Так, структура данных новой версии предполагает предварительное разбиение линий и полигонов прямоугольной сеткой с заданным шагом и последующее размещение полученных фрагментов в базе геоданных. Так как, при работе пользователя с системой основное время тратится на обращение систе-

мы к базе геоданных (поскольку при отрисовке карты необходимо постоянно получать координаты всех объектов, находящихся в прямоугольнике, ограниченном экраном), то данная схема позволяет существенно повысить скорость работы приложения.

1. Геоинформационные системы

1.1. Понятие и классификация ГИС

Геоинформационные системы (ГИС) — системы, предназначенные для сбора, хранения, анализа и графической визуализации пространственных данных и связанной с ними информации о представленных в ГИС объектах. Другими словами, это инструменты, позволяющие пользователям искать, анализировать и редактировать цифровые карты, а также дополнительную информацию об объектах, например высоту здания, адрес, количество жильцов. ГИС включают в себя возможности СУБД, редакторов растровой и век-

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 16-07-00089 и 18-07-20030, 18-37-10008.

торной графики и аналитических средств и применяются в картографии, геологии, метеорологии, землеустройстве, экологии, муниципальном управлении, транспорте, экономике, обороне [1].

По территориальному охвату различают глобальные, или планетарные ГИС, субконтинентальные ГИС, национальные ГИС, зачастую имеющие статус государственных, региональные ГИС, субрегиональные ГИС и локальные, или местные ГИС.

ГИС различаются предметной областью информационного моделирования, к примеру, городские ГИС, или муниципальные ГИС, природоохранные ГИС и т.п.; среди них особое наименование, как особо широко распространенные, получили земельные информационные системы. Проблемная ориентация ГИС определяется решаемыми в ней задачами (научными и прикладными), среди них инвентаризация ресурсов (в том числе кадастр), анализ, оценка, мониторинг, управление и поддержка принятия решений.

Интегрированные ГИС совмещают функциональные возможности ГИС и систем цифровой обработки изображений (материалов дистанционного зондирования) в единой интегрированной среде.

Полимасштабные, или масштабно-независимые ГИС основаны на множественных, или полимасштабных представлениях пространственных объектов, обеспечивая графическое или картографическое воспроизведение данных на любом из избранных уровней масштабного ряда на основе единственного набора данных с наибольшим пространственным разрешением. Пространственно-временные ГИС оперируют пространственно-временными данными.

Научные, технические, технологические и прикладные аспекты проектирования, создания и использования ГИС изучаются геоинформатикой. Геоинформатика — наука, технология и производственная деятельность по научному обоснованию, проектированию, созданию, эксплуатации и ис-

пользованию географических информационных систем, по разработке геоинформационных технологий, по приложению ГИС для практических и научных целей.

1.2. Принципы работы ГИС

ГИС хранит информацию о реальном мире в виде набора тематических слоев, которые объединены на основе географического положения. Этот простой, но очень гибкий подход оказывается полезным при решении разнообразных прикладных задач, таких как отслеживание передвижения транспортных средств и материалов, детальное отображение реальной обстановки и планируемых мероприятий, моделирование глобальной циркуляции атмосферы. Любая географическая информация содержит сведения о пространственном положении, будь то привязка к географическим или другим координатам, или ссылки на адрес, почтовый индекс, избирательный округ или округ переписи населения, идентификатор земельного или лесного участка, название дороги и т.п. При использовании подобных ссылок для автоматического определения местоположения или местоположений объекта (объектов) применяется процедура, называемая геокодированием. С ее помощью можно быстро определить и посмотреть на карте где находится интересующий нас объект или явление.

Векторная и растровая модели представления данных: ГИС может работать с двумя существенно отличающимися типами данных — векторными и растровыми. В векторной модели информация о точках, линиях и полигонах кодируется и хранится в виде набора координат X, Y . Местоположение точки (точечного объекта) описывается парой координат (X, Y) . Линейные объекты, такие как дороги, реки или трубопроводы, сохраняются как наборы координат X, Y . Полигональные объекты, типа речных водосборов, земельных участков или областей обслуживания, хранятся в виде замкнутого набора координат. Векторная модель особенно удобна для описа-

ния дискретных объектов и меньше подходит для описания непрерывно меняющихся свойств, таких как типы почв или доступность объектов. Растровая модель оптимальна для работы с непрерывными свойствами. Растровое изображение представляет собой набор значений для отдельных элементарных составляющих (ячеек), оно подобно отсканированной карте или картинке. Обе модели имеют свои преимущества и недостатки. Современные ГИС могут работать как с векторными, так и с растровыми моделями [2].

1.3. Основные задачи ГИС

ГИС общего назначения обычно выполняет пять основных процедур (задач) с данными: ввод, манипулирование, управление, запросы и анализ, визуализацию.

Ввод: Для использования в ГИС данные должны быть преобразованы в подходящий цифровой формат. Процесс преобразования данных с бумажных карт называется оцифровкой. В современных ГИС этот процесс может быть автоматизирован с применением сканерной технологии, что особенно важно при выполнении крупных проектов, либо, при небольшом объеме работ, данные можно вводить с помощью дигитайзера.

Манипулирование: Часто для выполнения конкретного проекта имеющиеся данные нужно дополнительно видоизменить в соответствии с требованиями вашей системы. Например, географическая информация может быть в различных масштабах (осевые линии улиц имеются в масштабе 1: 100 000, границы округов переписи населения — в масштабе 1: 50 000, а жилые объекты — в масштабе 1: 10 000). Для совместной обработки и визуализации все данные удобнее представить в едином масштабе. ГИС-технология предоставляет разные способы манипулирования пространственными данными и выделения данных, необходимых для конкретной задачи.

Управление: В небольших проектах географическая информация может храниться

в виде обычных файлов. Но при увеличении объема информации и росте числа пользователей для хранения, структурирования и управления данными эффективнее применять системы управления базами данных (СУБД), например в формате MS SQL. В ГИС наиболее удобно использовать реляционную структуру, при которой данные хранятся в табличной форме. При этом для связывания таблиц применяются общие поля.

Запросы и анализ: С помощью ГИС и географической информации можно решать как относительно простые задачи, такие как определение владельца данного земельного участка, определение расстояния между объектами, так и более сложные, требующие дополнительного анализа, запросы, например подбор места для строительства нового здания, определение типа почв под еловыми лесами, влияние строительства новой дороги на движение транспорта и так далее. Современные ГИС имеют множество мощных инструментов для анализа, среди которых наиболее значимы два: анализ близости и анализ наложения. Для проведения анализа близости объектов относительно друг друга в ГИС применяется процесс, называемый буферизацией. Он помогает ответить на вопросы типа: «Сколько домов находится в пределах 100 м от данного водоема?», «Сколько покупателей живет не далее 1 км от данного магазина?». Процесс наложения включает интеграцию данных, расположенных в разных тематических слоях. В простейшем случае это операция отображения, но при ряде аналитических операций данные из разных слоев объединяются физически.

Визуализация: Для многих типов пространственных операций конечным результатом является представление данных в виде карты или графика. Карта — это очень эффективный и информативный способ хранения, представления и передачи географической (имеющей пространственную привязку) информации. С помощью ГИС визуализация карт может быть легко до-

полнена отчетными документами, трехмерными изображениями, графиками и таблицами, фотографиями и другими средствами, например, мультимедийными [2].

1.4. Характеристика существующих ГИС

1.4.1. ArcGIS

Семейство программных продуктов разработано компанией ESRI, признанным лидером в сфере создания и продвижении ведущих Геоинформационных Систем, с учетом передовых тенденций развития информационных технологий и растущих требований многочисленных пользователей. Платформа ArcGIS 9 является оптимальным решением для построения корпоративной ГИС, фундамента информационной системы эффективного управления крупными государственными и коммерческими организациями.

На рис. 1 и 2 представлен пример интерфейса программного продукта семейства ArcGIS и пример создания трехмерного рельефа с помощью продукта ArcEditor.

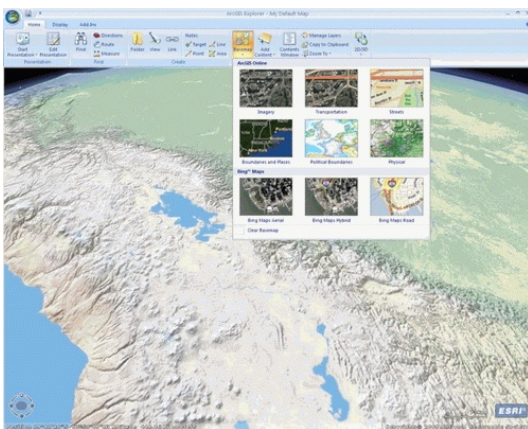


Рис. 1: Пример интерфейса ГИС ArcView

ArcGIS 9 построена на основе стандартов компьютерной отрасли, включая объектную архитектуру COM, .NET, Java, XML, SOAP, что обеспечивает поддержку общепринятых стандартов, гибкость предлагаемых решений, широкие возможности взаимодействия. Одной из основных особенностей данного семейства продуктов является наличие базы геоданных.

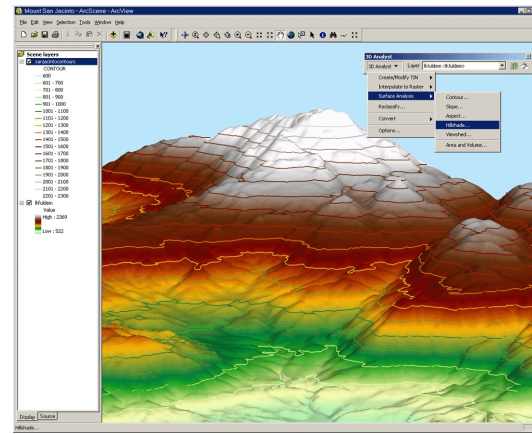


Рис. 2: Создание трехмерного рельефа с помощью ArcEditor

База геоданных — это созданная компанией ESRI модель, определяющая структуру и правила хранения различных видов данных — векторных и растровых, адресных точек, данных геодезических измерений, и многих других. Уникальная технология позволяет эффективно хранить разнородные данные и с легкостью использовать их в сложных проектах и системах. В базе геоданных пользователи могут задавать правила и отношения внутри хранилища, которые определяют поведение пространственно взаимосвязанных географических объектов и объектных классов и обеспечивают целостность данных (то, что обычно называют топологией). База геоданных позволяет проводить редактирование, как в многопользовательском режиме, так и в автономном, с возможностью синхронизации версий.

Геообработка — это анализ географической информации, одна из основных функций ГИС. ArcGIS 9 содержит более 450 инструментов: для проведения анализа, конвертации, управления данными, геокодирования, динамической сегментации, картографии, работы с растрами; от оверлейных операций, построения буферных зон, инструментов для выявления пространственных закономерностей и управления данными до расширенных возможностей: обработки растров, методов интерполяции и оценки качества данных, зональной фильтрации, многофакторного анализа, растровой алгебры,

построения и проверки топологии, построения графических схем.

К особенностям визуализации относятся:

- Производство качественных картографических продуктов со всеми необходимыми элементами зарамочного оформления, использованием прозрачности, собственных или уже готовых стандартных условных знаков, штриховок, градуированных символов, картограмм и диаграмм.
- Автоматическая генерация схем сетевых объектов, представление данных, изменяющихся во времени, а также возможность 3D визуализации, расширяющая области применения ГИС.

В ArcGIS можно быстро создать реалистичную виртуальную 3D сцену на основе пространственных данных, как локального уровня, так и в масштабе всей Земли, с использованием цифровых моделей рельефа, космических и аэроснимков, любых векторных данных и фотореалистичных моделей объектов. Работать с трехмерными объектами можно так же, как и со стандартными ГИС-слоями — делать выборки, получать атрибутивную информацию, оформлять сцены в соответствии с поставленными задачами [3].

Пакет программных продуктов ArcGis включает в себя как настольные, так и серверные ГИС.

Примерами настольных ГИС являются:

ArcGIS ArcView — базовый продукт семейства ArcGIS, полнофункциональная ГИС с набором мощных инструментов для создания, управления, анализа и визуализации пространственных данных.

ArcGIS ArcEditor — сочетает функциональность ArcView с возможностями создания и моделирования баз геоданных (БГД). Уникальный механизм обеспечивает поддержку целостности и многопользовательского редактирования БГД, управление версиями, построение топологии и геометрических сетей.

ArcGIS ArcInfo — расширяет функциональность вышеперечисленных продуктов (ArcView, ArcEditor) с набором мощных инструментов для пространственного анализа и геообработки данных.

Примерами серверных ГИС являются:

ArcGIS Server предназначен для создания корпоративной ГИС с неограниченным числом полнофункциональных рабочих мест: клиентом может быть как настольное, так и веб-приложение. ArcGIS Server предоставляет инструментарий для создания веб-приложений, веб-служб и других корпоративных приложений, работающих под управлением стандартных .NET и J2EE веб-серверов, обеспечивает централизованное управление географическими ресурсами: картами, службами геокодирования и программными объектами, задействованными в приложениях.

ArcIMS — продукт для публикации пространственных данных и картографической продукции в интранет/интернет с возможностью геокодирования, анализа и поиска данных по различным критериям. Служит основой для создания порталных решений, работает под управлением стандартных веб-серверов.

ArcSDE — обеспечивает хранение пространственных данных в СУБД (Oracle, Microsoft SQL Server, IBM DB2 и Informix). Для представления и хранения информации используется объектно-реляционная модель — база геоданных, позволяющая описывать не только геометрию объектов, но и их поведение, правила, взаимосвязи с другими классами объектов и объектами базы геоданных. Поддерживается работа с различными версиями данных, длительные сеансы редактирования и автономное редактирование. ArcSDE обеспечивает интеграцию ArcGIS с другими ГИС системами.

1.4.2. Google Earth

Google Earth — проект компании Google, в рамках которого в сети Интернет были размещены спутниковые фотографии всей

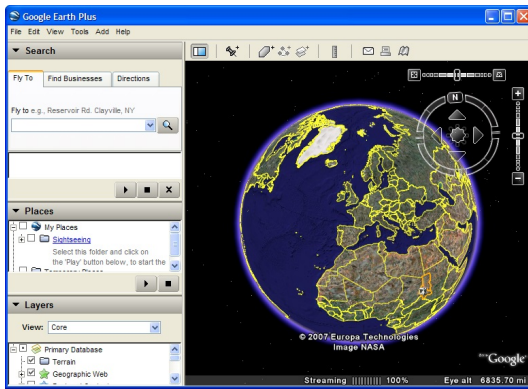


Рис. 3: Интерфейс программы Google Earth



Рис. 4: Пример изображения населенного пункта

земной поверхности. Фотографии некоторых регионов имеют очень высокое разрешение. В отличие от других аналогичных сервисов, показывающих спутниковые снимки в обычном браузере (например, Google Maps), в данном сервисе используется специальная, загружаемая на компьютер пользователя клиентская программа Google Earth. Такой подход обеспечивает дополнительные возможности, трудно реализуемые с помощью веб-интерфейса.

На рис. 3 показан пример интерфейса Google Earth, на рис. 4 и 5 видно, насколько детально Google Earth передает изображение местности.

Основные возможности программы [4]:

1. Google Earth автоматически подкачивает из интернета необходимые пользователю изображения и другие данные, сохраняет их в памяти компьютера и на жёстком диске для дальнейшего использования. Скачанные данные сохраняют-



Рис. 5: Пример изображения береговой линии

ся на диске, и при последующих запусках программы закачиваются только новые данные, что позволяет существенно экономить трафик.

2. Для визуализации изображения используется трёхмерная модель всего земного шара (с учётом высоты над уровнем моря), которая отображается на экране при помощи интерфейсов DirectX или OpenGL. Именно в трёхмерности ландшафтов поверхности Земли и состоит главное отличие программы Google Earth от её предшественника Google Maps. Пользователь может легко перемещаться в любую точку планеты, управляя положением «виртуальной камеры».
3. Практически вся поверхность суши покрыта изображениями, полученными от компании DigitalGlobe, и имеющими разрешение 15 м на пиксель. Есть отдельные участки поверхности (как правило, покрывающие столицы и некоторые крупные города большинства стран мира), имеющие более подробное разрешение. Например, Москва снята с разрешением 0,6 м/пиксель, а многие города США — с разрешением 0,15 м/пиксель. Данные ландшафта имеют разрешение порядка 100 м.
4. Кроме того, имеется огромное количество дополнительных данных, которые можно подключить по желанию пользователя. Например, названия населённых пунктов, водоёмов, аэропортов, до-

роги, ж/д, и другая информация. Кроме того, для многих городов имеется более подробная информация — названия улиц, магазины, заправки, гостиницы, и т. д.

5. Пользователи могут создавать свои собственные метки и накладывать свои изображения поверх спутниковых (это могут быть карты, или более детальные снимки, полученные из других источников).
6. В программе есть слой «3D Здания», с трёхмерными моделями, добавляемыми разработчиками или самими пользователями, посредством сервиса 3D Warehouse. В городах России можно найти модели некоторых значимых памятников архитектуры.
7. Имеется упрощённая Java-версия программы для сотовых телефонов.
8. Реализована функция измерения расстояний между объектами.

1.4.3. ГИС ПАНОРАМА

Геоинформационная система ПАНОРАМА предназначена для решения на базе персональных компьютеров следующих задач:

- Создание и обновление электронных карт по исходным картографическим материалам или по материалам аэрофотосъёмки.
- Ведение пользовательской базы данных на отображаемую местность.
- Решение прикладных задач с использованием картографической и атрибутивной баз данных.
- Создание автоматизированных систем управления процессами, использующими данные о местности.

От аналогичных продуктов ГИС ПАНОРАМА отличается высокой скоростью визуализации (обработки) картографических данных в условных знаках, принятых для топографических, обзорно-географических, кадастровых и других видов карт.

ГИС ПАНОРАМА разработана специалистами Топографической Службы Воору-

женных Сил Российской Федерации. На базе системы в подразделениях Топографической Службы производится промышленное создание цифровых карт местности различных масштабов и назначения.

Основой ГИС ПАНОРАМА является специализированная система управления базами данных электронных карт, предназначенная для создания и обновления векторных, растровых и матричных карт, а также для разработки прикладных задач в 32-х разрядных операционных системах.

База данных электронных карт имеет иерархическую структуру. На нижнем уровне хранится информация об отдельных объектах карты. Объекты могут объединяться в группы, слои и листы карт. Совокупность листов карт одного масштаба и вида составляет район работ — отдельную базу данных электронных карт. Описание отдельного объекта состоит из метрических данных (координат на местности), семантических данных (свойств объекта), текстовых справочных данных, иллюстративных графических данных и других данных, включая уникальный номер объекта, через который осуществляется логическая связь с внешними реляционными СУБД.

Система «ПАНОРАМА» обеспечивает отображение района работ, состоящего из нескольких листов, как единого целого, что облегчает решение прикладных задач на больших территориях. Отдельные листы района работ могут быть в любой момент обновлены, отредактированы без необходимости дополнительных действий над остальными листами района работ.

Основными функциями системы управления базами данных электронных карт являются:

- создание иерархической структуры базы данных электронных карт, имеющей уровни: район работ, листы карт, слои объектов, объекты местности;
- редактирование содержимого базы данных с использованием графического интерфейса пользователя: создание ново-

- го уровня, удаление, обновление, копирование, восстановление;
- поддержка различных проекций и систем координат;
- визуализация содержимого базы данных в условных знаках, принятых для топографических, обзорно-географических, кадастровых, и других видов карт, быстрый скроллинг изображения; изменение состава отображаемых объектов;
- вывод на внешние устройства печати изображения электронной карты в принятых условных знаках, поддержка векторных и растровых устройств печати, цветных и черно-белых. Состав объектов и масштаб карты могут изменяться, выводимое на печать изображение может отображаться в режиме WYSIWYG;
- поддержка стандартных систем классификации и кодирования объектов и их характеристик, в соответствии с требованиями Роскартографии, ВТС МО РФ и других федеральных служб;
- поддержка пользовательских условных знаков, слоев, объектов и их характеристик. Для отображения условных знаков могут применяться графические примитивы, не поддерживаемые в GDI системы Windows (пунктирные линии произвольной толщины и длины, заполнение площадного объекта точечными условными знаками размером больше 8 на 8 и т. д.), обеспечивается корректный вывод этих условных знаков на печатающие устройства с высокой скоростью;
- выполнение запросов на поиск объектов с заданными характеристиками;
- выполнение расчетных операций: определение площади, длины, периметра, направления и т. д.;
- выполнение справочных запросов по топологии объектов: поиск точек пересечения, примыкания, поиск объектов, расположенных внутри или вне заданного объекта, на заданном расстоянии от объекта;
- построение зон вокруг объектов, построение пересечений объектов, отображе-

ние результатов выполнения запросов на фоне электронной карты;

- совместная обработка векторных, растровых и матричных данных о местности, создание, отображение, вывод на печать растрово-векторных электронных карт (трансформированные фрагменты раstra, имеющие произвольную границу для каждого из них, отображаются и обрабатываются как единое целое с постепенным наполнением векторными объектами).

1.4.4. ГИС MapInfo

MapInfo — географическая информационная система, предназначенная для сбора, хранения, отображения, редактирования и анализа пространственных данных. Первая версия ГИС MapInfo Professional была разработана в 1987 году компанией MapInfo Corp., и стала одной из самых популярных ГИС в мире. Сейчас MapInfo Professional используется в 130 странах мира, переведена на 20 языков, включая русский, и установлена в десятках тысяч организаций. В России благодаря простоте освоения, богатым функциональным возможностям и разумной стоимости, MapInfo Professional стала самой массовой геоинформационной системой [5].

ГИС MapInfo — высокоэффективное средство для визуализации и анализа пространственных данных. Сферы применения ГИС MapInfo: бизнес и наука, образование и управление, социологические, демографические и политические исследования, промышленность и экология, транспорт и нефтегазовая индустрия, землепользование и кадастр, службы коммунального хозяйства и быстрого реагирования, армия и органы правопорядка, а также многие другие отрасли хозяйства. Пример интерфейса ГИС MapInfo приведен на рис. 6.

С помощью ГИС MapInfo решаются следующие задачи:

- поставщики услуг сотовой связи используют MapInfo для улучшения качества

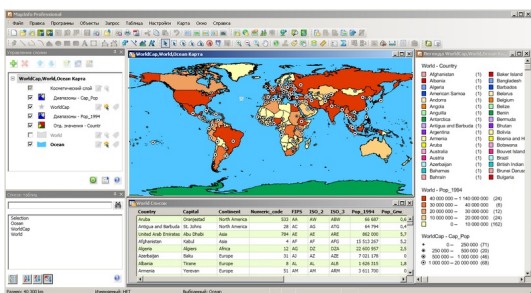


Рис. 6: Пример интерфейса ГИС MapInfo

услуг и расширения зоны обслуживания;

- транспортные компании используют MapInfo для планирования и оптимизации маршрутов доставки грузов;
- торговые компании с помощью MapInfo могут следить за динамикой продаж, проводить маркетинговый анализ, планировать размещение торговых точек;
- страховые компании используют MapInfo для оценки степеней риска для данной территории;
- правоохранительные органы используют MapInfo для анализа оперативной обстановки и обеспечения общественной безопасности;
- органы государственной власти применяют MapInfo для территориального планирования и ведения земельного и других кадастров;
- геологи и маркшейдеры используют MapInfo при разведке и добыче полезных ископаемых.

Поддерживаемые форматы данных.

Помимо собственных форматов, MapInfo работает без конвертации с графическими данными в форматах ArcView Shape File, ESRI ArcSDE, ESRI Geodatabase (mdb), ARC/INFO E00, AutoCAD DXF/DWG, Intergraph/MicroStation Design DGN, SDTS, VPF и табличными данными в форматах Access, Excel, Lotus 1-2-3, xBASE и ASCII. Универсальный транслятор MapInfo позволяет осуществлять импорт и экспорт данных в другие ГИС и САПР системы (ESRI Shape File, AutoCAD DXF/DWG, Intergraph/MicroStation Design DGN, AtlasGIS, ARC/INFO E00).

Работа с растровыми изображениями MapInfo имеет возможность работы с данными в растровых форматах GIF, JPEG, TIFF, GEO TIFF, PCX, BMP, TGA, BIL и др., включая новейшие форматы сжатого раstra — ECW, MrSID, JPEG2000.

Поиск информации и создание запросов.

Встроенный язык запросов SQL, благодаря географическому расширению, позволяет осуществлять выборки объектов с учетом их пространственных отношений. MapInfo имеет функции поиска объекта или группы объектов по различным признакам, а также их сочетаниям.

Интересной особенностью программы MapInfo является использование векторизатора MapEDIT. Программы MapEDIT/MapEDIT PRO предназначены для создания и редактирования цифровых векторных карт, а также других графических данных, по растровым изображениям, получаемым в результате сканирования карт, планов, фотоснимков с бумажных, пластиковых и других твердых носителей.

Создаваемые с помощью MapEDIT/MapEDIT PRO цифровые карты отвечают всем современным требованиям, предъявляемым существующими геоинформационными системами для цифровых векторных карт.

Данная технология, используя многочисленные элементы автоматизации трудоемкого процесса оцифровки картографических материалов, позволяет многократно сократить время и повысить качество создания электронных цифровых карт по сравнению с традиционной технологией, основанной на использовании планшетных дигитайзеров.

При работе с программой оператор одновременно наблюдает на экране компьютера, как растровое изображение исходных материалов, так и создаваемое (редактируемое) векторное изображение карты, что предоставляет возможности контроля качества векторизации. Мощная инструментальная

палитра MapEDIT/MapEDIT PRO содержит все необходимые инструменты, от ручных процедур до полностью автоматических, для редактирования растровых и векторных слоев электронной цифровой карты.

Используемые растровые изображения:

- программы MapEDIT/MapEDIT PRO могут работать с растровыми изображениями более чем 30 форматов (PCX, BMP, TIFF, GIF, JPEG и т.д.) произвольной цветности и практически неограниченного размера;
- возможно использование нескольких пересекающихся или не пересекающихся растров различных масштабов для создания по ним единой векторной картой.

Обработка растровых изображений:

- набор процедур изменения растрового изображения: удаление/замена цвета, уменьшение размеров полей заданного цвета, сглаживание, удаление небольших точек (пятен) заданного цвета, рисование и т.д.;
- сшивка отдельных растровых фрагментов в единое растровое изображение (возможность подготовки растровых изображений на сканерах небольшого формата);
- трансформация (выравнивание) растра с использованием произвольной сети картографических реперов;
- сохранение преобразованных растровых изображений или их фрагментов в любом из стандартных форматов (PCX, BMP, TIFF, GIF, JPEG и других).

Фотограмметрическая обработка аэро- и космических снимков:

- специальные преобразования растровых изображений для устранения геометрических и фотометрических искажений, обусловленных кривизной Земли, неровностями рельефа местности, погрешностями съемочной аппаратуры и других,

- характерных для снимков, а так же преобразования, связанные с переводом в заданную картографическую проекцию;
- возможность обработки двух видов снимков: — обычного проекционного типа (аэро- и космические снимки) и сканерного типа (космические снимки);
- использование снимков с частично и неточно заданными параметрами.

Создание, редактирование и контроль векторных данных:

- задание структуры векторной карты (разделение объектов по слоям и типам);
- задание структуры базы атрибутивных данных, наполнение базы непосредственно на этапе создания карты;
- ручные процедуры редактирования объектов векторной карты;
- автоматизированные процедуры трассировки линий различного типа и оконтуривания полигонов;
- использование априорных данных о форме объектов (ортогональность, параллельность и т.п.) в ручных и автоматизированных процедурах векторизации;
- одновременная поддержка двух моделей представления цифровой карты (линейно-узловой и объектовой);
- полный контроль (автоматический и интерактивный) топологической корректности карт, обнаружение и устранение ошибок топологии по произвольно выбранным векторным слоям или их комбинациям; построение полигональных покрытий;
- трансформация (выравнивание) векторной карты по произвольному набору картографических реперов;
- сводка («сшивка») векторных карт, вводимых по отдельным планшетам, для получения единой векторной карты региона.

Связь с ГИС: экспорт/импорт графических данных вместе со связанными атрибутивными данными в большинство су-

ществующих ГИС и САПР в реальных или условных координатах, в виде полигональных покрытий или в виде объектов MapEDIT/ MapEDIT PRO.

2. Программно-функциональная система «Геоанализ»

2.1. Описание «ПФС-Геоанализ»

Программно-функциональная система геоанализа («ПФС – Геоанализ») — комплекс визуального анализа для решения оперативно-аналитических задач с использованием геонформационных технологий.

«ПФС – Геоанализ» включает в себя [6]:

1. Программно-функциональный комплекс мониторинга мобильных объектов («ПФК – Мониторинг»), обеспечивающий, загрузку и отображение плановых маршрутов движения, получение и обработку фактических данных от средств навигации, контроль передвижения и анализ — сопоставления плановых и фактических данных, подготовку отчетных документов.
 2. Программно-функциональный комплекс — Геоанализ («ПФК – Геоанализ»), реализованный в архитектуре:
 - «толстого» клиента с расширенными возможностями по связыванию атрибутивных данных с картографической основой, отображению и редактированию графических и атрибутивных данных, сортировке, поиску, фильтрации атрибутивных данных, подготовке и выполнению аналитических запросов, подготовке и сохранению отчетов в формате Word и Excel;
 - «тонкого» клиента с возможностью доступа к картографической информации с использованием Internet, обеспечивающий отображение картографических данных и результатов экспресс-анализа, поиск объектов или событий, связанных с картой, выполнение аналитических за-
- просов, подготовку отчетов в формате HTML.

Основными задачами «ПФС – Геоанализ» являются:

1. Визуализация, ввод и редактирование картографической и алфавитно-цифровой информации, с возможностью динамического пересчета координат при отображении картографической основы.
2. Импорт растровых изображений, содержащих дополнительную информацию (космические снимки, планы, схемы и т.п.) и отображение их на цифровой карте наряду с объектами картографической основы.
3. Геокодирование — автоматическое размещение на картографической основе объектов или событий на основе анализа адресной информации.
4. Импорт/экспорт геоданных в/из форматы dxf/dbf, sxf, mif/mid, shapefile с возможностью загрузки и совместного отображения данных, представленных в различных системах координат и проекциях.
5. Информационно-справочный режим для получения информации по объектам и событиям, связанным с картографической основой; с возможностью сохранения состояния фрагментов карты для быстрого перехода от одного состояния к другому и выполнения поисковых операций для активных слоев.
6. Пространственный анализ структурированной информации о территории, объектах и событиях для получения отчетов в виде тематических многоцветных карт.
7. Создание буферных зон и решение информационно-справочных, аналитических задач на основе набора таких зон.
8. Планирование и расчет маршрутов, на основе данных транспортной сети.
9. Мониторинг мобильных объектов, сопоставление полученных данных с плановыми заданиями.

Основными блоками, входящими в состав «ПФС – Геоанализ» являются:

1. Блок «Ввод и редактирование данных», обеспечивающий ввод и редактирование алфавитно-цифровой и графической информации, относящейся к объектам или событиям, размещенным на картографической основе.
2. Блок «Геокодирование», обеспечивающий загрузку, обработку и автоматическое размещение на картографической основе объектов или событий из «внешних» баз данных на основе анализа адресной информации.
3. Блок «Информационно-справочное обеспечение», предназначенный для поиска и визуализации в табличном виде и на картографической основе данных по интересующим объектам или событиям, в том числе отображение космических снимков совместно с картографической основой.
4. Блок «Пространственный экспресс-анализ», предназначенный для оперативного анализа объектов или событий, а также охватываемой ими территории для получения тематически раскрашенных карт.
5. Блок «Планирование перемещений» особо опасных грузов по территории, предназначенный для прокладки маршрутов, расчета их протяженности и временного графика передвижения транспортных средств.
6. Блок «Мониторинг», поддерживающий сбор, обработку и анализ фактических данных о мобильных объектах, сопоставление полученных данных с плановыми заданиями.

2.2. Переход к новой версии системы

В настоящее время, в связи с постоянно растущим объемом картографических данных, разработанный ранее программный комплекс перестал удовлетворять жестким требованиям пользователей к быстродействию системы. В связи с этим возникла необходимость разработки новой версии си-

стемы, которая могла бы работать значительно быстрее предыдущей.

При работе пользователя с системой основное время тратится на обращение системы к базе геоданных (так как при отрисовке карты необходимо постоянно получать координаты всех объектов, находящихся в прямоугольнике, ограниченном экраном). Поэтому для ускорения исполнения SQL — запросов в новой версии применяется разбиение исходных данных прямоугольной сеткой, для ускорения процесса выбора объектов, попавших в конкретную ячейку. Кроме того, осуществляется генерализация данных, что значительно повышает быстродействие системы за счет сокращения объема данных.

2.3. Понятие и задачи генерализации

Генерализация — процессы отбора и обобщения содержания при составлении географических карт. Основной целью генерализации является сохранение и выделение на карте основных, типических черт и характерных особенностей изображаемых явлений в соответствии с назначением данной карты, её тематикой и возможностями масштаба.

Генерализация проявляется: в отборе объектов (то есть в ограничении содержания карты необходимыми объектами и в исключении прочих), в продуманном упрощении контуров, то есть плановых очертаний объектов — линейных и площадных (при котором сохраняются, а иногда даже усиливаются особенности контура, характерные для данного объекта), в обобщении количественных характеристик объектов [7].

Наиболее часто встречающимися способами генерализации являются:

1. Отбор — удаление мелких объектов в соответствии с целевым масштабом.
2. Удаление «лишних» точек, лежащих в пределах одного и того же пикселя на экране для данного масштаба.

3. Преобразование площадных объектов, ширина которых менее одного пикселя, в линейные.
4. Удаление узких перемычек и сглаживание контуров объектов.
5. Объединение близко расположенных объектов с учетом целевого масштаба.

3. Структура ПО «Геоконвертор»

3.1. Технологическое описание процесса конвертации

В процессе подготовки картографического проекта можно выделить четыре технологических этапа функционально связанных между собой:

1. **Импорт исходных параметров картографического проекта** На данном этапе пользователь производит импорт информации об иерархии слоев картографического проекта, индивидуальных настройках каждого слоя, приоритете отрисовки слоев и так далее.
2. **Настройка пользователем параметров проекта** На данном этапе пользователь производит настройку многих параметров конвертации, как отдельных слоев, так и проекта в целом, наиболее важными из которых являются следующие: система координат проекта, масштабы видимости слоев, применяемые алгоритмы генерализации.
3. **Автоматическая конвертация данных в соответствии с заданными параметрами** Автоматическая конвертация данных происходит без участия пользователя и может занимать длительное время. В процессе конвертации формируется отчет о текущем состоянии конвертации (количество обработанных слоев, выявленные ошибки и так далее).
4. **Контроль качества сконвертированного проекта** Контроль качества сконвертированного проекта выполняется пользователем после завершения конвертации и включает в себя анализ ошибок, указанных в сформированном

отчете, а также визуальный контроль сконвертированного картографического проекта.

Процесс конвертации схематически можно представить следующим образом (рис. 7).

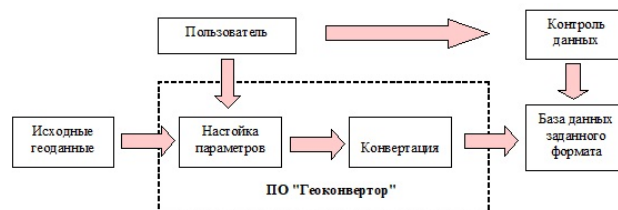


Рис. 7: Схема процесса конвертации

Учитывая описанные выше особенности процесса конвертации, при разработке ПО «Геоконвертор» была применена следующая модульная схема построения приложения.

3.2. Модульная схема ПО «Геоконвертор»

Согласно приведенной выше схеме процесса конвертации была выбрана модульная структура ПО «Геоконвертор». Основными функциональными модулями ПО «Геоконвертор» являются:

- настраиваемый интерфейс пользователя;
- модуль импорта исходных данных;
- модуль обработки данных на основе алгоритмов генерализации;
- модуль экспорта данных в базу.

3.2.1. Пользовательский интерфейс

Основной задачей пользовательского интерфейса является настройка параметров конвертации картографического проекта. Интерфейс приложения позволяет эффективно управлять глобальными настройками проекта (список масштабов проекта, целевая система координат, параметры конвертации атрибутивных данных), управлять иерархией проекта (добавлять/удалять слои, объединять их в группы и так далее), управлять надстройками конкретного слоя или группы слоев.

Пример интерфейса приложения в части настроек глобальных параметров проекта приведен на рис. 8, а параметров картографического слоя на рис. 9.

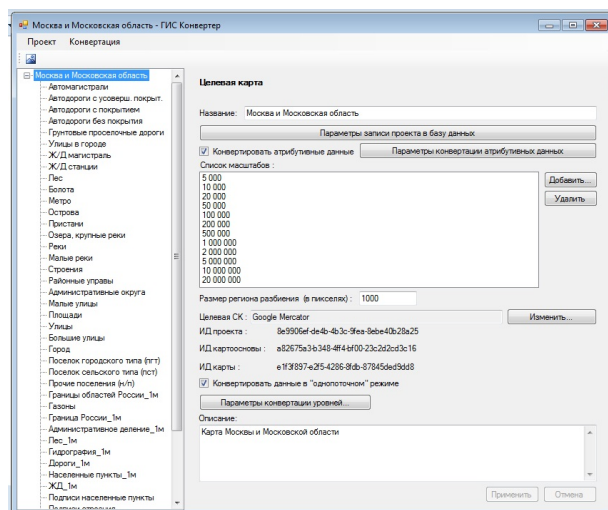


Рис. 8: Пример пользовательского интерфейса ПО «ГеоКонвертер» (настройка глобальных параметров проекта)

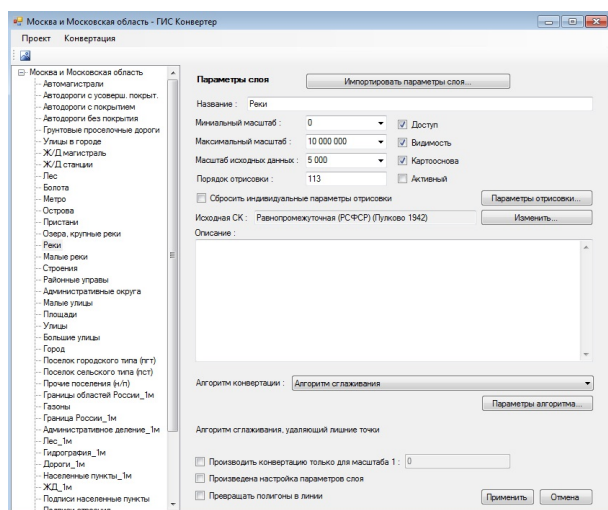


Рис. 9: Пример пользовательского интерфейса ПО «ГеоКонвертер» (настройка параметров картографического слоя)

3.2.2. Импорт данных

Основными функциями модуля импорта данных являются импорт настроек картографического проекта (в состав настроек входят иерархия слоев проекта и индивидуальные настройки каждого слоя), а также

импорт геоданных из источников в различных форматах.

Импорт геоданных возможен из следующих форматов: формат старой версии ПФС «Геоанализ» (база данных MS-SQL), формат старой версии ПФС «Геоанализ» (файлы .typ/.dat), формат данных ГИС MapInfo (файлы .dat/.tab/.map/.id).

3.2.3 Обработка данных

Данный модуль является основным модулем ПО «ГеоКонвертер». В его основе лежат классы, осуществляющие процесс автоматической конвертации. Взаимодействие модулей ПО «ГеоКонвертер» представлено на рис. 10.

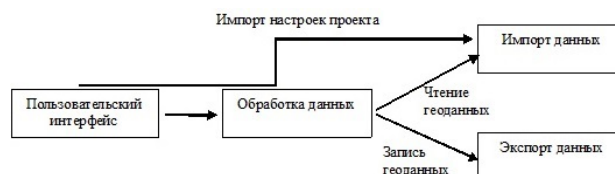


Рис. 10: Взаимодействие модулей ПО «ГеоКонвертер»

В процессе конвертации проекта модуль обрабатывает последовательно слои проекта, считывая исходные данные для каждого из слоев, применяя к ним алгоритм генерализации, выбранный пользователем для конкретного слоя, и экспортируя полученные данные с помощью модуля вывода. Указанные алгоритмы обработки геоданных описаны в следующей главе.

3.2.4. Экспорт данных

Основной задачей данного модуля является запись иерархии картографического проекта и обработанных геоданных в базу MS SQL заданного формата совместимого с новой версией ПФС «Геоанализ». Важной частью данного модуля является класс, обеспечивающий разбиение линейных и полигональных объектов на квадраты со стороны, зависящей от масштаба, для которого производится конвертация. Данная операция производится непосредственно перед записью объектов в базу, после применения алгоритмов генерализации.

4. Алгоритмы обработки данных

В ПО «Геоконвертор» разработано несколько алгоритмов обработки данных, что позволяющих эффективно обрабатывать различные типы геоданных. Сложностью вычислений называют оценку, позволяющую определить как быстро растёт трудоёмкость алгоритма с увеличением объема данных. Далее приведено подробное описание алгоритмов.

4.1. «Тривиальный» алгоритм

Тип входных данных. Гео-объекты любого типа.

Основные параметры. Целевой масштаб карты.

Описание алгоритма. Алгоритм выполняет копирование данных без дополнительной обработки, то есть оставляет данные как они есть.

Данный алгоритм обычно применяется для конвертации слоев растровой графики, сформированных заранее названий и подписей, а также при конвертации слоев линейных и полигональных объектов, не требующих генерализации (например, поэтажные планы зданий).

Техническая реализация. Данный алгоритм получает исходные данные различных типов и транслирует их в модуль экспорта без каких-либо изменений. Все объекты картографического слоя обрабатываются последовательно.

Оценка временной сложности алгоритма. Данный алгоритм имеет линейную временную сложность, в общем случае оценка его сложности (N), где N – суммарное количество отрезков, импортируемых из исходных данных

4.2. Алгоритм сглаживания

Тип входных данных. Картографические слои, содержащие объекты полигонального или линейного типа

Основные параметры. Целевой масштаб карты, минимальная площадь полигональных объектов, минимальное расстояние между соседними точками и максимальное отклонение от прямой при сглаживании.

Описание алгоритма. Основной идеей данного алгоритма является исключение избыточных (с учетом целевого масштаба) точек контура объекта, что позволяет существенно снизить объем данных без потери качества отображения карты мелкого масштаба. Основными этапами алгоритма являются:

- удаление гео-объектов площадь, которых меньше заданного порогового значения;
- удаление соседних точек контура объекта, расстояние между которыми меньше заданного значения;
- сглаживание объектов — удаление точек контура объекта по критерию параметра максимального отклонения от исходного контура при сглаживании.

Данный алгоритм является основным алгоритмом, используемым при генерализации картографические слои, содержащих объекты полигонального или линейного типа. Алгоритм активно применяется при обработке таких слоев как границы стран и регионов, гидрография, леса, автомобильные дороги и так далее.

Техническая реализация. Основные этапы алгоритма:

1. Удаление «мелких» объектов — на данном этапе все полигональные объекты обрабатываются последовательно, те из них, которые имеют площадь меньше заданного порогового значения, полностью удаляются.
2. Удаление соседних точек контура — на данном этапе все контуры гео-объектов обрабатываются последовательно, для каждой точки объекта последовательно удаляются соседние точки, расстояние до которых меньше расстояния, заданного для данного масштаба. Далее

процедура повторяется для следующей в порядке обхода контура точки.

3. Сглаживание контуров — на данном этапе все контуры гео-объектов обрабатываются последовательно. Основная идея алгоритма представлена на рис. 11. Рассмотрим пару несмежных точек контура (точки A и C). Лежащая между ними точка B проверяется на критерий сглаживания (расстояние до прямой, проходящей через точки A и C). В случае если расстояние до прямой не превышает заданную величину, точка B исключается из контура. Данная процедура повторяется для всего контура.

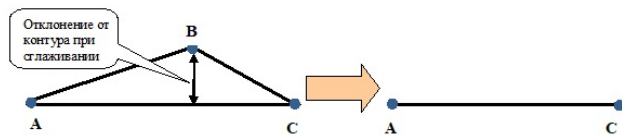


Рис. 11: Основной принцип сглаживания контура

Оценка временной сложности алгоритма.

Примем следующие константы:

- Количество объектов полигонального и линейного типа, импортируемых из исходных данных.

N – Среднее количество точек в контуре гео-объекта.

Проведем оценку сложности по шагам (таблица 1):

Таблица 1: Временная сложность этапов алгоритма сглаживания

Этапы алгоритма	Временная сложность
Удаление «мелких» объектов	$O(KN)$
Удаление соседних точек контура	$O(KN^2)$
Сглаживание контуров	$O(KN^2)$

Определим итоговую сложность алгоритма сглаживания, которая равняется сумме полученных значений, так как все шаги алгоритма выполняются последовательно один за другим.

Итоговая сложность: $O(KN) + O(KN^2) + O(KN^2)$.

Это выражение эквивалентно $O(KN^2)$.

На практике самым медленным является третий этап. Однако эмпирическая сложность существенно меньше теоретической и обычно не превышает $O(KN \cdot \log(N))$.

4.3. Алгоритм обработки точечных объектов

Тип входных данных. Точечные объекты (объекты, отображающиеся на карте в виде «иконки» заданного размера и расположенной в точке с заданными координатами)

Основные параметры. Целевой масштаб карты, размер отображаемой на карте «иконки» (в пикселях)

Описание алгоритма. Алгоритм выполняет перенос данных точечных гео-объектов с учетом технических особенностей формата представления данных. Данный алгоритм обычно применяется для конвертации различных слоев, содержащих точечные объекты (станции метро, железнодорожные станции, населенные пункты и так далее).

Техническая реализация. Основной технической особенностью реализации данного алгоритма является необходимость дублирования объекта в случае попадания области занимаемой «иконкой» на целевом масштабе в несколько квадратов сетки разбиения. Основная идея алгоритма проиллюстрирована на рис. 12.

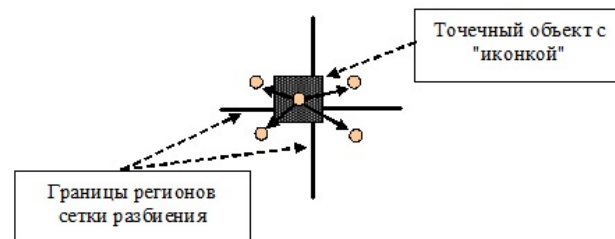


Рис. 12: Дублирование точечных объектов, захватывающих несколько регионов разбиения

Оценка временной сложности алгоритма. Данный алгоритм имеет линейную временную сложность, в общем случае оценка его сложности $O(N)$, где N – суммарное количество точечных объектов, импортируемых из исходных данных.

4.4. Алгоритм генерализации кварталов

Тип входных данных. Картографический слой «Строения» (объекты полигонального типа), картографические слои гео-объектов, которые могут являться границами кварталов (автомагистрали, улицы, площади, парки, гидрографические объекты).

Основные параметры. Целевой масштаб карты, минимальная плотность строений в квартале, минимальная площадь квартала.

Описание алгоритма.

В связи с большим количеством и высокой плотностью строений в городах и других населенных пунктах, отображение каждого из них в отдельности становится нецелесообразным на картах масштаба более мелкого, чем 1:20 000. Таким образом, становится актуальной следующая задача: объединение стоящих рядом строений в «кварталы» с последующим отображением на карте слоя кварталов вместо слоя строений. Данную задачу решает алгоритм генерализации кварталов.

Под кварталом будем понимать многоугольник, стороны которого совпадают с различными гео-объектами (улицами, границами парков, площадей, реками и т.д.), причем отношение суммарной площади строений, лежащих внутри данного многоугольника, к площади самого многоугольника должно быть не меньше определенного заранее значения [8, 9].

На рис. 13 схематично представлена основная задача алгоритма — объединение стоящих рядом строений (на рисунке выделены желтым цветом) в кварталы (выделены зеленым).



Рис. 13: Ожидаемый результат работы алгоритма

Техническая реализация.

Основная идея решения поставленной задачи состоит в разбиении исследуемой области на простые многоугольники, сторонами которых являются отрезки, содержащиеся в объектах, которые могут служить границами кварталов (автомагистралях, улицах, площадях, парках, гидрографических объектах), проверке каждого из них на наличие в нем достаточного количества сооружений и последующем объединении соседних кварталов, в соответствии с картографическим требованиям для каждого целевого масштаба карты. Основные этапы алгоритма:

1. Генерация «сетки» перекрестков.
2. Построение точек пересечения отрезков, которые в дальнейшем станут вершинами искомым кварталов.
3. Построение многоугольников по получившейся «сетке». На данном этапе вся обрабатываемая территория разбивается на простые многоугольники, вершинами которых являются построенные перекрестки, а сторонами — отрезки, по которым строилась «сеть».
4. Проверка получившихся многоугольников на принадлежность к кварталам на основании плотности застройки. Для каждого построенного многоугольника определим список строений, лежащих внутри него, и вычислим их суммарную площадь, затем вычислим площадь самого многоугольника и вычислим их отношение. В случае, если отношение не меньше заданного параметра, будем считать данный многоугольник кварталом.
5. Слияние соседних кварталов. Для каждого целевого масштаба на основании «базового» слоя кварталов, полученного

в пункте 3, будем проводить объединение соседних (то есть имеющих общую сторону) кварталов, до тех пор, пока все кварталы не станут удовлетворять картографическим требованиям для данного масштаба. Причем, генерацию слов будет производить «снизу вверх», то есть начиная от самого крупного масштаба и заканчивая самым мелким, заменяя на каждом шаге «базовый» слой на вновь сформированный.

6. Обработка получившихся кварталов с учетом конкретного масштаба. На данном этапе производится параллельный сдвиг сторон кварталов (необходим для отображений ненулевой ширины улиц), удаление «перемычек» — то есть очень узких частей кварталов, и сглаживание слишком острых углов, получившихся в результате сдвига.

Таблица 2: Временная сложность этапов алгоритма генерализации кварталов

Этапы алгоритма	Временная сложность
Генерация «сетки» перекрестков	$O(N) + O(N(\log(N))) + O(N^2) + O(N^2) + O(NQ)$
Построение многоугольников	$O(N^2(Q/M)) + O(N) = O(N^2(Q/M))$
Проверка многоугольника на принадлежность к кварталам	$O(BM)$
Слияние кварталов	$O(N^2(\log(N))/M) + NM(\log(N))$
Обработка получившихся кварталов с учетом конкретного масштаба	$O(N^2/M)$

Подробное описание технической реализации алгоритма генерализации кварталов приведено в [10].

Оценка временной сложности алгоритма.

Примем следующие константы:

N — Суммарное количество отрезков, импортируемых из исходных данных.

Q — Среднее число отрезков, пересекающихся в одной точке.

B — Суммарное количество строений в исследуемой области.

M — Количество многоугольников, разбивающих исследуемую область.

Проведем оценку сложности по шагам (таблица 2):

Определим итоговую сложность алгоритма генерализации кварталов, которая равняется сумме полученных значений, так как все шаги алгоритма выполняются последовательно один за другим.

Итоговая сложность:

$$O(N) + O(N^2) + O(N^2(Q/M)) + O(BM) + O(N^2(\log(N))/M + NM(\log(N))) + O(N^2/M).$$

Это выражение эквивалентно $O(N^2 + MB + NM(\log(N)))$, так как множители Q/M и $\log(N)/M$ на практике много меньше единицы и ими можно пренебречь. На практике, при относительно малом количестве отрезков, самым медленным этапом является проверка полигонов на принадлежность к кварталам, однако с ростом числа отрезков самыми медленными становятся этапы генерации и слияния многоугольников — этот результат полностью согласуется с теорией (это очевидно из формулы для суммарной сложности алгоритма).

Детальная оценка временной сложности каждого этапа алгоритма приведена в [10].

4.5. Алгоритм обработки линий уровня

Линия уровня или изолиния — условное обозначение на карте, чертеже, схеме или графике, представляющее собой линию, в каждой точке которой измеряемая величина (высота над уровнем моря) сохраняет одинаковое значение.

Тип входных данных. Картографический слой полигональных объектов (линии уровня), отметки о высоте каждой из линий уровня.

Основные параметры. Целевой масштаб карты, искомая плотность линий уровня.

Описание алгоритма. Участки местности с резкими перепадами высот на картах мелкого масштаба имеют высокую плотность линий уровня, что препятствует визуальному восприятию данных фрагментов карты. Для решения данной проблемы будем применять алгоритм «прореживания» линий уровня, то есть удаления некоторых из них таким образом, чтобы плотность линий не превышала заданной пользователем величины.

Для отдельных участков местности рассчитаем шаг сетки высот, который позволит достичь требуемой плотности линий и удалим все линии уровня, высота которых не принадлежит узлам сетки высот. Данный принцип представлен на рис. 14.

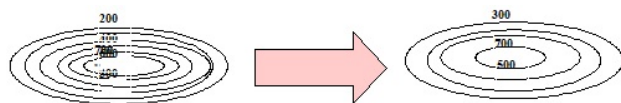


Рис. 14: Принцип работы алгоритма обработки линий уровня

Данный алгоритм применяется для обработки слоев, содержащих высотные линии уровня, изобаты и других слоев, содержащих информацию о рельефе поверхности.

Техническая реализация. Первым шагом вычисляется шаг сетки высот, который позволит достичь требуемой плотности линий. Далее исходные линии уровня обрабатываются последовательно, в случае если высота линии не принадлежит узлам построенной сетки высот, данная линия удаляется.

Оценка временной сложности алгоритма.

Данный алгоритм имеет линейную временную сложность, в общем случае оценка его сложности (N), где N – суммарное количество линий уровня, содержащихся во входных данных.

4.6. Алгоритм формирования подписей

Основной задачей данного алгоритма является пространственное расположение на-

званий и подписей объектов на карте соответствующего масштаба с целью улучшения визуального представления данных.

Тип входных данных. Картографические слои объектов, для которых необходимо сгенерировать подписи, атрибутивная информация (текст подписи для каждого объекта)

Основные параметры. Целевой масштаб карты, приоритет подписей слоев, тип генерируемых подписей

Описание алгоритма. Данный алгоритм позволяет производить генерацию и расположение подписей объектов четырех разных типов:

- улицы;
- линейная гидрография;
- площадная гидрография;
- строения и населенные пункты.

Отбор подписей производится с учетом принципа «конкуренции» слоев: если несколько подписей гео-объектов различных слоев не могут быть изображены на карте совместно (без наложения друг на друга), то выбирается та из них, слой которой имеет наивысший приоритет. Данный принцип проиллюстрирован на рис. 15.

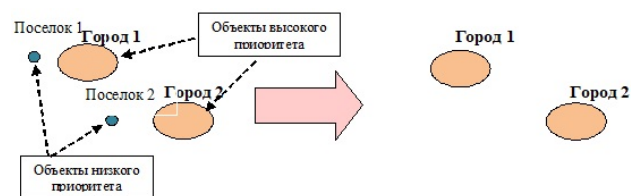


Рис. 15: Иллюстрация принципа «конкуренции» слоев по приоритету

Данный алгоритм применяется для генерации подписей различных гео-объектов (названия населенных пунктов, названия улиц, номера домов, названия рек, озер и так далее).

Техническая реализация.

Алгоритм можно разбить на два этапа:

1. генерация подписей для каждого объекта;

2. выбор подписей в соответствии с принципом «конкуренции» слоев по приоритету.

На первом этапе способ генерации надписей зависит от типа объекта.

- 1) Подписи названий улиц — основной идеей расположения подписей названий улиц является расположение надписи вдоль отрезков достаточной длины или групп почти параллельных отрезков (угол между которыми лежит в заданных пределах). Данный подход проиллюстрирован на рис. 16.



Рис. 16: Принцип расположения названий улиц

- 2) Подписи объектов линейной гидрографии — основной способ расположения подписей линейных гидрографических объектов аналогичен способу расположения названий улиц. Отличие заключается в необходимости контроля плотности подписей одного объекта. Данный метод проиллюстрирован на рис. 17.

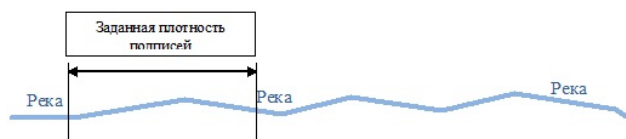


Рис. 17: Способ расположения подписей линейного гидрографического объекта

- 3) Подписи объектов площадной гидрографии — основной способ расположения надписей площадной гидрографии — расположение надписей горизонтально как можно ближе к середине обрамляющего гео-объект прямоугольника. Данный метод проиллюстрирован на рис. 18. В случае если расположить надпись таким образом нельзя (нет горизонтального отрезка соответствующей длины, лежащего полностью внутри объекта), то производится проверка пар не соседних точек контура

на предмет возможности расположения подписи наклонно вдоль соединяющего их отрезка (отрезок должен полностью лежать внутри объекта и иметь достаточную длину).



Рис. 18: Способ расположения подписей площадного гидрографического объекта

- 4) Подписи строений и населенных пунктов — алгоритм расположения подписей строений и населенных пунктов состоит из двух шагов: сначала подпись размещается горизонтально в центре прямоугольника, обрамляющего объект (см. рис. 19 слева). Затем, если подпись выходит за границы объекта, производится сдвиг в точку, полученную следующим образом:

- из центра обрамляющего объект прямоугольника (точка C) проводим лучи, касающиеся объекта (точки A и B на рис. 19);
- проводим биссектрису угла C и находим точки M и D ее пересечения с объектом;
- середина отрезка MD является искомой точкой.

Оценка временной сложности алгоритма.

Примем следующие константы:

K — Количество объектов, для которых нужно сгенерировать подписи.

N — Среднее количество точек в контуре гео-объекта.

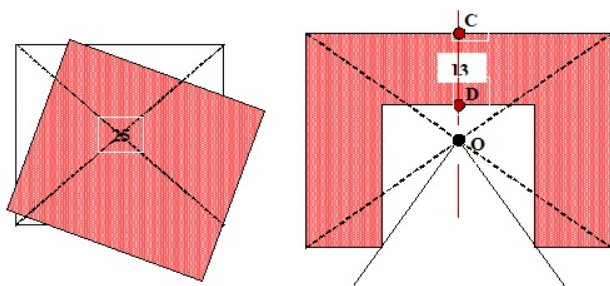


Рис. 19: Способ расположения подписей строений и населенных пунктов

Приведем оценку сложности каждой из вариаций алгоритма (таблица 3):

Таблица 3: Временная сложность алгоритма формирования подписи

Тип алгоритма	Временная сложность
Подписи названий улиц	$O(KN) + O(K^2)$
Подписи объектов линейной гидрографии	$O(KN) + O(K^2)$
Подписи объектов площадной гидрографии	$O(KN^2) + O(K^2)$
Подписи строений и населенных пунктов	$O(KN) + O(K^2)$

5. Результаты и выводы

Основным результатом данной работы является теоретическая разработка и практическая реализация ПО «ГЕОКОНВЕРТОР» в составе программно-функционального комплекса «Геоанализ».

Разработанное программное обеспечение имеет модульную структуру и включает в себя следующие блоки:

- настраиваемый интерфейс пользователя;
- модуль импорта исходных данных;
- модуль обработки данных на основе алгоритмов генерализации;
- модуль экспорта данных в базу.

Интерфейс разработанного приложения позволяет пользователю управлять глобальными настройками проекта (список масштабов проекта, целевая система координат, параметры конвертации атрибутивных данных), управлять иерархией проекта (добавлять/удалять слои, объединять их в

группы и так далее), управлять настройками конкретного слоя или группы слоев. Основными функциями модуля импорта данных являются импорт настроек картографического проекта (в состав настроек входят иерархия слоев проекта и индивидуальные настройки каждого слоя), а также импорт геоданных из источников различных форматов.

Модуль экспорта данных позволяет производить запись иерархии картографического проекта и обработанных геоданных в базу MS SQL заданного формата совместимого с новой версией ПФС «Геоанализ», обеспечивая разбиение линейных и полигональных объектов на полигоны сетки разбиения.

Наиболее важной частью ПО «Геоконвертор» является модуль обработки геоданных. Данный модуль включает в себя следующие основные алгоритмы обработки данных:

- «Тривиальный» алгоритм;
- Алгоритм сглаживания;
- Алгоритм обработки точечных объектов;
- Алгоритм генерализации кварталов;
- Алгоритм обработки линий уровня;
- Алгоритм формирования подписей.

Далее приведем несколько примеров карт, сконвертированных с помощью ПО «Геоконвертор».

На рис. 20 приведем фрагмент карты Москвы и Московской области. На данном фрагменте подписи названий населенных пунктов сформированы с помощью «Алгоритма формирования подписей». Также граница области обработана с помощью «Алгоритм сглаживания».

На рис. 21 приведен фрагмент той же карты в более крупном масштабе 1:200 000. На данном фрагменте виден результат работы «Алгоритма формирования подписей» для населенных пунктов и крупных гидрографических объектов, а также «Алгоритма сглаживания».

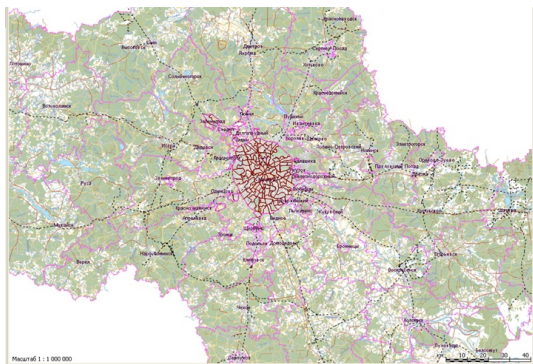


Рис. 20: Фрагмент карты Москвы и Московской области в масштабе 1:1 000 000



Рис. 21: Фрагмент карты Москвы и Московской области в масштабе 1:200 000

Результат работы «Алгоритма формирования подписей» для улиц в городе и номеров домов можно видеть на рис. 22 и 23, на которых приведен фрагмент карты в масштабах 1:50 000 (рис. 22) и 1:10 000 (рис. 23).



Рис. 22: Фрагмент карты в масштабе 1:50 000

Отдельно продемонстрируем результат работы «Алгоритм генерализации кварталов». На рис. 24 приведен фрагмент карты северо-восточного округа города Моск-

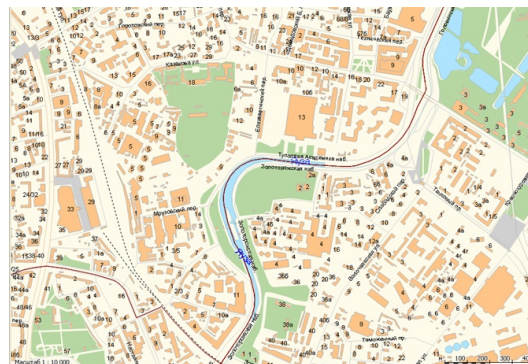


Рис. 23: Фрагмент карты в масштабе 1:10 000

вы, на котором изображены все строения, лежащие в иллюстрируемой области.

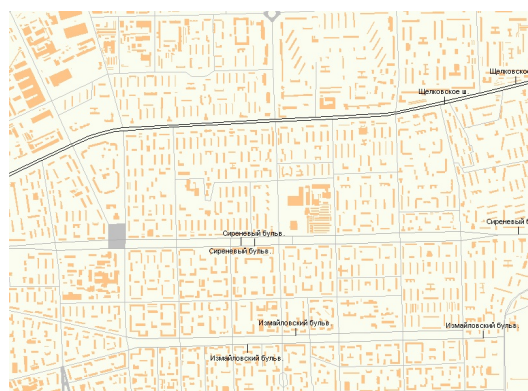


Рис. 24: Фрагмент карты с изображением большого количества строений

Приведем результат работы алгоритма генерализации кварталов для целевых масштабов 1:50 000 (рис. 25) и 1:200 000 (рис. 26).

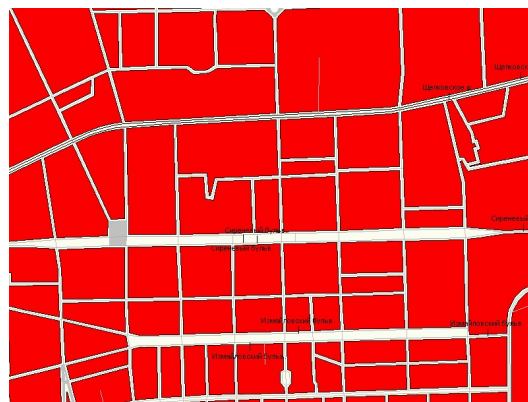


Рис. 25: Результат работы алгоритма генерализации кварталов для целевого масштаба 1:50 000

Как видно из рис. 25 и 26 количество геообъектов в слое «Кварталы» существенно

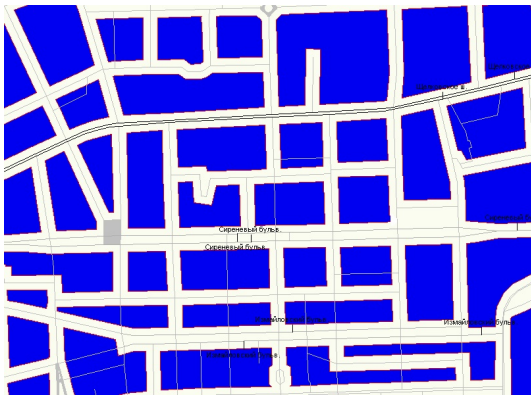


Рис. 26: Результат работы алгоритма генерализации кварталов для целевого масштаба 1:200 000

меньше количества строений на участке исходной карты.

Литература

- [1] Журкин И.Г., Шайтура С.В. Геоинформационные системы. — М., «КУДИЦ-ПРЕСС», 2009.
- [2] http://www.dataplus.ru/Industries/100_GIS/GIS.htm.
- [3] <http://www.dataplus.ru/Soft/ESRI/ArcGIS/ArcGIS.htm>.
- [4] http://ru.wikipedia.org/wiki/Google_Earth.
- [5] <http://www.esti-map.ru>.
- [6] Босов А.В., Петров М.Е., Судейкин М.И. Опыт решения оперативно-аналитических задач с использованием геоинформационных технологий. // Системы высокой доступности. 2007. Т. 3, №4, с. 24–32.
- [7] Салищев К.А., Филиппов Ю.В. Основы генерализации на общегеографических картах мелкого масштаба. — Картография, М., 1971.
- [8] Наставление по составлению и подготовке к изданию топографических карт масштабов 1:10 000, 1:25 000, 1:50 000, 1:100 000. М.: Редакционно-издательский отдел ВТС, 1980.
- [9] Руководство по картографическим и картоиздательским работам. Часть 2. Составление и подготовка к изданию топографических карт масштабов 1:200 000, 1:500 000. М.: Редакционно-издательский отдел ВТС, 1980.
- [10] Афанасьев А.А. Разработка и реализация алгоритмов генерализации в составе геоинформационной системы «Геоанализ», выпускная квалификационная работа бакалавра, МФТИ, 2010.

Анализ особенностей АСУ ТП АЭС при решении задач демонстрации функциональных возможностей и обучения пользователей*

В. Ф. Захарушкин
zakharushkin@mail.ru

Протвино, Институт физико-технической информатики

В работе рассматриваются особенности АСУ технологическими процессами функционирования атомных электростанций (АЭС). Обосновывается специфика АЭС в отличие от других типов электростанций. Приводится состав подсистем на уровне блока АЭС и направления развития АСУ ТП, требующие решения задач новыми методами визуализации обрабатываемой информации и обучения персонала.

Ключевые слова: АСУ ТП, атомная электростанция (АЭС), состав подсистем, визуализация данных, обучение пользователей.

Analysis of the features of the APCS of NPPs in solving the tasks of demonstrating functional capabilities and training users*

V. F. Zakharushkin

Protvino, Institute of Computing for Physics and Technology

The paper discusses the features of ACS technological processes of nuclear power plants (NPPs) functioning. The specifics of nuclear power plants are justified, in contrast to other types of power plants. The composition of the subsystems at the NPP unit level and the directions of the automated process control system development are presented, which require the solution of problems by new methods of visualizing the processed information and training the personnel.

Keywords: APCS, nuclear power plant (NPP), subsystems composition, data visualization, user training.

1. Введение

Амбициозные цели по созданию АО «РАСУ» и назначению его интегратором нового бизнеса ГК «Росатом» — «Автоматизированные системы управления технологическими процессами», поставили перед данной организацией задачи обеспечения развития и повышения конкурентоспособности отечественных АСУ ТП на международном и российском рынках автоматизированных систем управления.

В современном мире, когда информационные технологии развиваются опережающими темпами, и пронизывают большинство других технологий, АСУ ТП является одним из ведущих направлений развития материального производства. Важнейшим фактором конкурентоспособности проектных решений в настоящее время являются такие технологические процессы, как автоматизированный марке-

тинг (выставочно-демонстрационная деятельность), поддержка эксплуатации и обучение пользователей. Указанные выше технологические процессы были исторически в СССР, а затем и в России менее развиты, чем в передовых странах Запада.

Вторым важным аспектом внедрения современных информационных технологий, влияющих на конкурентоспособность отечественных проектных решений в ядерных технологиях, являются средства обучения и информационно-справочной поддержки пользователей АСУ ТП АЭС. Эффективные средства обучения снижают трудоёмкость эксплуатации и повышают устойчивость системы управления АЭС к потенциальным сбоям и отказам.

В настоящее время предстоит в короткие сроки не только устранить указанное отставание отечественных разработок от потенциальных конкурентов, но и выйти в указанных выше направлениях на передовые рубежи.

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 17-07-00833, 18-07-20030.

2. Анализ особенностей (специфики) АСУ ТП

2.1. Особенности АСУ ТП АЭС

Автоматизированная система управления технологическими процессами атомной электростанции (АСУ ТП АЭС) имеет свои особенности, которые могут повлиять на выбор средств решения поставленных задач.

Автоматизированная система управления технологическим процессом (АСУ ТП) — группа решений технических и программных средств, предназначенных для автоматизации управления технологическим оборудованием на промышленных предприятиях [1]. Может иметь связь с более общей автоматизированной системой управления предприятием (АСУП) или корпоративной информационной системой (КИС) управления организацией.

Под АСУ ТП обычно понимается целостное решение, обеспечивающее автоматизацию основных операций технологического процесса на производстве в целом или каком-то его участке, выпускающем относительно завершённое изделие.

Соответственно, под АСУ ТП АЭС понимается автоматизированная информационная система, предназначенная для управления программно-аппаратными средствами функционирования отдельного блока АЭС или станции в целом

Отличительной особенностью АСУ ТП АЭС от АСУ ТП большинства других направлений промышленности являются следующие:

- повышенные требования к функциям надёжности работы оборудования;
- повышенные требования к безопасности проведения работ;
- тяжёлые последствия аварийных ситуаций;
- необходимость соответствия требованиям как отечественных, так и международных организаций (МАГАТЭ);

- большое количество сопрягаемых подсистем, применяемых технологий и проектных решений.

Если первые пять тезисов для атомной энергетики являются достаточно очевидными, то последний требует определённых пояснений.

Подтверждением последнего утверждения является то, что в АСУ ТП АЭС контролируется около 350 000 показателей, а служит первичными данными для вывода на пультах дежурной смены 200 000 параметров. Эти цифры характеризуют сложность системы и подтверждают необходимость применения современных информационных технологий для решения поставленных задач.

2.2. Краткое описание АСУ ТП АЭС

Для описания АСУ ТП АЭС кратко представим что собой представляет АЭС. На рис. 1 представлена принципиальная схема блока АЭС, как составной части станции.

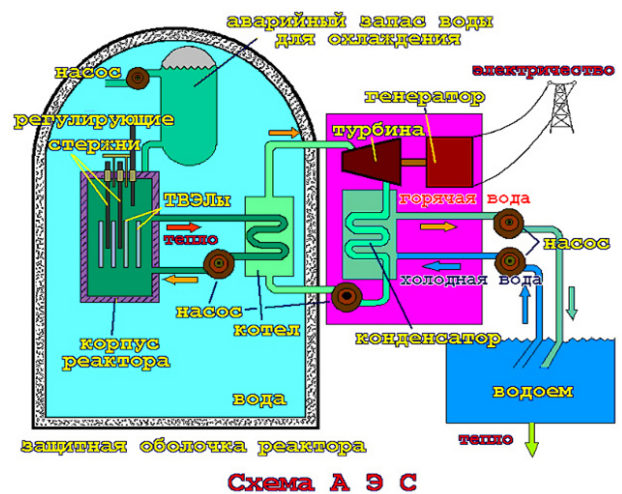


Рис. 1: Принципиальная схема функционирования блока АЭС

Из представленной схемы видно, что АЭС отличается от тепловой электростанции наличием ядерного реактора и составом систем, обеспечивающим его работу.

2.3. Автоматизированная система контроля и управления технологическими процессами

предназначена для реализации следующих функций [3]:

- управления всеми технологическими объектами станции в условиях, определяемых проектом АЭС;
- контроля технологических объектов управления (ТОУ) и автоматизированного ведения эксплуатационных режимов АЭС, защиты оборудования, а также автоматического регулирования параметров ТОУ;
- диагностики основного оборудования;
- информационного обеспечения персонала во всех эксплуатационных режимах АЭС.

Для выполнения основного назначения АЭС — выработки электроэнергии, в АСУ ТП присутствует ряд систем нормальной эксплуатации (часть из которых относятся к важным для безопасности), которые предназначены для контроля и управления технологическими, электротехническими, вентиляционными и пр. системами, связанными с ведением технологического процесса по выработке электроэнергии.

Для выполнения всех возлагаемых на АСУ ТП задач, включая обеспечение безопасности, в ее состав входят:

- датчики;
- импульсные трубные проводки за отсечными устройствами или коренными вентилями;
- стенды и элементы для установки датчиков;
- средства оперативно-диспетчерского и местного дистанционного управления, блокировок, защит, сигнализации и авторегулирования;
- программно-технические средства обработки, хранения и передачи информации;
- средства отображения и регистрации информации (вторичные приборы, дисплеи, табло, индикаторы и т.п.);

- местные посты и посты центрального управления с оперативно-командными элементами;
- средства приема и выдачи информационных и управляющих дискретных сигналов на исполнительные устройства;
- кабели, кабельные коммуникации и проходки;
- средства электропитания средств АСУ ТП от шин РУСН, включая устройства пониженного напряжения;
- средства автоматизации, поставляемые комплектно с технологическим оборудованием;
- средства радиационного контроля;
- низковольтные комплектные устройства электропитания арматуры;
- аппаратура, стенды, оборудование и инструменты, обеспечивающие обслуживание, ремонт и метрологическое обеспечение средств ТС АСУ ТП.

Оборудование и технические процессы на атомной электростанции обладают рядом особенностей, в значительной степени определяющих требования к АСУ ТП.

К этим особенностям относятся:

- работа оборудования в условиях радиационных нагрузок, высоких давлений и температур, которая является результатом как быстропротекающих, так и инерционных ядерно-физических и тепловых процессов;
- недоступность значительной части оборудования во время работы на мощности;
- необходимость обеспечения радиационной и ядерной безопасности, как при нормальной эксплуатации, так и при аварийных ситуациях и авариях;
- необходимость обеспечения пожаро-, взрыво- и электробезопасности, надежности и экономичности работы АЭС;
- сложность и многообразие основного и вспомогательного оборудования: большое количество запорной и регулирующей арматуры, механизмов, устройств, агрегатов, большое число и значительное разнообразие измеряемых па-

раметров (температура, давление, расход, уровень, механические перемещения, электрические и радиационные измерения, химконтроль и др.);

- разнообразие целей и требований, предъявляемых к управлению технологическими процессами;
- высокие стоимости простоев энергоблока;
- климатические условия;
- сейсмические условия;
- требования по экологии (охрана окружающей среды, воздушного бассейна, почвы, воды).

Все эти особенности делают атомную электростанцию сложным объектом управления, требующим высокой степени автоматизации оборудования и централизации управления, применения современных средств вычислительной техники, высоконадежной и эффективной системы управления, позволяющей небольшому количеству обслуживающего персонала осуществлять управление технологическими процессами.

2.4. Определение уровня автоматизации технологических задач АЭС

При определении уровня автоматизации технологических задач АЭС с использованием функционально-группового управления принимаются следующие общие положения:

- Поддержание (по мере возможности) процесса выработки электроэнергии при нарушении нормальной эксплуатации;
- Уровень автоматизации должен обеспечивать выполнение требований по глубоко эшелонированной защите, направленных на предотвращение нарушений нормальной эксплуатации, предотвращение проектных аварий, предотвращение запроектных аварий, управления авариями с целью снижения их последствий;
- Человек-оператор рассматривается как главное, центральное звено управления,

выполняющее наиболее ответственные задачи по управлению энергоблоком в режимах нормальной эксплуатации. При этом человеческий фактор играет важную роль в обеспечении безопасности станции и сохранении барьеров безопасности. Для взаимодействия с АСУ ТП, в состав которой не входит оператор, предусмотрен ЧМИ, предоставляющий возможность получения оператором необходимой информации и формирования команд.

Основными критериями при определении уровня автоматизации технологических задач АЭС являются:

- снижение возможных ошибок оперативного персонала при выполнении рутинных и ответственных задач управления;
- экономическая эффективность АЭС и затраты на обслуживание АСУ ТП;
- обеспечение безопасного управления энергоблоком.

Требования к объёму автоматизации процессов обработки данных энергоблока АЭС представлены в таблице 1.

Конструкция и характеристики применяемого на АЭС технологического оборудования должны обеспечивать возможность реализации принятого в концепции уровня автоматизации технологических процессов.

Возможности по реализации требуемого уровня автоматизации управления энергоблоком АЭС со стороны АСУ ТП обеспечены следующими решениями:

- выбором необходимых программно-технических средств для обеспечения возможности реализации заданных функций автоматизированного и автоматического контроля и управления и учетом возможности расширения этих функций при необходимости;
- выбором необходимой для реализации поставленных задач контроля и управления структуры комплекса технических средств АСУ ТП, обеспечивающей возможность подсистемам АСУ ТП обмениваться информацией и командами

Таблица 1: Объем автоматизации энергоблока

Технологическая область энергоблока	Количество				
	Точек контроля	Электроприводной запорной арматуры	Насосов, вентиляторов, электронагревателей	Регулирующей арматуры	ФГУ
Реакторное отделение, системы безопасности (один канал)	850	650	130	15	—
Реакторное отделение — системы НЭ, НЭ ВВ, вспомогательные	1450	380	100	50	30
Турбинное отделение — системы НЭ, НЭ ВВ, вспомогательные	1650	320	100	90	30
СВО	550	600	130	60	43
Вентиляция, для обеспечения работы систем безопасности	—	40	30	—	—
Огнезадерживающие, клапана НЭ	—	500	—	—	—
Огнезадерживающие, клапана в каналах систем безопасности	—	90	—	—	—
Итого	4500	2580	490	215	103

управления для реализации поставленных задач;

- выбором современных, модернизированных с учетом опыта эксплуатации, имеющих достаточные показатели надежности измерительных преобразователей (датчиков) и исполнительных механизмов; применением развитых средств диагностики (датчиков, исполнительных механизмов и собственно программно-технических средств).

2.5. Структура АСУ ТП

Структура АСУ ТП имеет иерархический принцип построения в соответствии с разделением энергоблока, как объекта управления, на технологические функциональные части.

АСУ ТП предназначена для обеспечения:

- контроля и управления основными и вспомогательными технологическими процессами на энергоблоке АЭС;
- обеспечения безопасности во всех режимах работы, включая аварийные ситуации и аварии;
- повышения экономичности работы АЭС в условиях нормальной эксплуатации.

АСУ ТП функционирует непрерывно в течение всего срока службы энергоблока.

АСУ ТП предусматривает следующие режимы работы: первоначальный запуск, штатное функционирование, вывод компонентов АСУ ТП из работы и ввод в работу, реконфигурацию при неисправностях.

В режиме первоначального запуска осуществляется включение питания и загрузка ПО. При включении системы производится самодиагностика и исключается выдача ложных команд и ложной сигнализации.

В режиме штатного функционирования АСУ ТП выполняет все функции, включая самодиагностику, а также осуществляет непрерывный и периодический контроль исправности технических и целостности программных средств АСУ ТП.

В режиме вывода из эксплуатации элементов АСУ ТП для ремонта или техобслуживания, как в объекте управления, так и в АСУ ТП — ложные команды и информация не формируется. При вводе в работу отдельных устройств выполняется диагностика и обеспечивается задержка на время обновления информации и входа датчиков в

диапазон приемлемой точности. Факты нарушений в работе АСУ ТП отображаются, записываются и хранятся на соответствующих средствах АСУ ТП.

При этом передача в систему недостоверного результата измерения без указания о его недостоверности исключена.

При неисправностях, возникающих в резервированных структурах АСУ ТП, осуществляется автоматическая реконфигурация.

2.6. Подсистемы АСУ ТП энергоблока

АСУ ТП энергоблока включает следующие подсистемы:

- Система верхнего блочного уровня (СВБУ);
- Системы управления и защиты реактора и УСБТ (СУЗ-УСБТ);
- Системы контроля и управления оборудованием нормальной эксплуатации (СКУ НЭ), включая:
 - 1) ПТК реакторного отделения (ПТК СКУ РО);
 - 2) ПТК контроля и управления турбинного отделения (ПТК СКУ ТО);
 - 3) ПТК системы контроля и управления оборудованием специальной водоочистки (ПТК СКУ СВО);
 - 4) ПТК системы контроля и управления вентиляционным оборудованием (ПТК СКУ В);
 - 5) ПТК системы контроля и управления вспомогательным оборудованием турбогенератора (ПТК СКУ ТГ);
 - 6) ЭЧСЗ — шкафы собственных защит турбины и генератора (разрабатывается по отдельному от ПТК СКУ НЭ ЧТЗ);
 - 7) ЭЧСР — электронная часть системы регулирования турбины (разрабатывается по отдельному от ПТК СКУ НЭ ЧТЗ);
 - 8) датчики технологических параметров систем НЭ.
- Автоматизированная система радиационного контроля (АСРК);
- Система контроля, управления и диагностики РУ (СКУД), включающая в себя систему внутриреакторного контроля (СВРК), а также диагностические системы: система контроля вибрации (СКВ), система обнаружения течи теплоносителя первого контура (СОТТ), система обнаружения свободных предметов (СОСП), система комплексного анализа (СКА), система автоматизированного контроля остаточного ресурса оборудования (САКОР) РУ;
- Система обнаружения течи теплоносителя второго контура (СОТТ-2);
- Система контроля и управления противопожарной защитой (СКУ ПЗ);
- Система контроля и управления электротехническим оборудованием НЭ (СКУ ЭЧ);
- Система контроля и управления водно-химического режима (СКУ ВХР) — реализуется как распределенная задача СКУД, СКУ НЭ, СВБУ и АРМ лабораторий ВХР, в составе которой обеспечивается анализ, прогнозирование и выдача рекомендаций персоналу по ведению ВХР;
- Система регистрации важных параметров эксплуатации (СРВПЭ);
- МПУ — местные пункты управления, включая систему контроля гидроамортизаторов (СКГА).

Подсистемы АСУ ТП имеют следующие назначения:

СВБУ — система, выполняющая общеблочные функции и объединяющая все подсистемы АСУ ТП в единую систему управления технологическими процессами энергоблока.

СУЗ-УСБТ — управляющая система, предназначенная для контроля и управления реактором, выполнения нормального и аварийного останова реактора по теплотехническим, нейтронно-физическим параметрам, поддержания его в подкритическом состоянии. В ней выделена система УСБТ — управляющая система безопасности по тех-

нологическим параметрам, предназначенная для реализации команд иницилирующей части СУЗ-УСБТ, дистанционного управления с целью формирования сигналов на запуск технологических систем безопасности (активных), контроля и управления защитных, обеспечивающих и локализирующих системы безопасности, а также реализации защиты и блокировок НЭ.

АСРК — система, предназначенная для контроля радиационных параметров, их обработки и выдачи обработанной информации.

СКУД — система, предназначенная для контроля состояния активной зоны и реакторной установки в целом, формирования сигналов ограничения мощности реактора, управления полем энерговыделения по активной зоне, а также диагностики состояния основного оборудования РУ.

СКУ ПЗ — система, предназначенная для автоматического обнаружения возникновения пожара, сигнализации и запуска систем пожаротушения, а также контроля и управления ими, вентсистемами и системами дымоудаления в процессе ликвидации пожара.

СКУ НЭ — система контроля и управления оборудованием НЭ, включая:

- ПТК СКУ РО — система, предназначенная для выполнения функций защит и блокировок, дистанционного управления, технологической сигнализации и авторегулирования, применительно к технологическим системам нормальной эксплуатации реакторного отделения.
- ПТК СКУ ТО — система, предназначенная для выполнения задач защит и блокировок, дистанционного управления, авторегулирования и технологической сигнализации применительно к технологическим системам турбинного отделения, за исключением задач, решаемых в ЭЧСР и ПТК, ТГ и технологических защит собственно турбины и генератора (ЭЧСЗ).
- ПТК СКУ ТГ — система, предназначенная для температурного контроля гене-

ратора, а также контроля и управления технологическими системами генератора.

- ПТК СКУ СВО система, предназначенная для автоматизации задач контроля и управления системами спецводоочистки I и II контуров энергоблока.
- ПТК СКУ В — система, предназначенная для автоматизации задач контроля и управления системами вентиляции энергоблока, включая контроль за обеспечивающими системами вентиляции.

ЭЧСР — система, предназначенная для контроля и регулирования параметров турбоустановки, а также контроля тепломеханических величин, характеризующих работу турбины.

ЭЧСЗ — шкафы (ПТК) собственных технологических защит турбины и генератора;

Датчики технологических параметров соответствующих технологических подсистем ЭБ АЭС, являющиеся интерфейсом с технологическим объектом контроля и управления в части контроля.

СРВПЭ — система регистрации важных параметров при нарушениях нормальной эксплуатации в предаварийных и аварийных ситуациях.

СКУ ВХР — система/задача/ контроля и управления водно-химическим режимом контура; указанная задача реализуется в ПТК ТО.

АСВД — система вибродиагностики основного оборудования (турбогенератор, ГЦН, циркуляционные насосы и питательные насосы).

2.7. Структура АСУ ТП энергоблока

АСУ ТП энергоблока представляет собой децентрализованную и пространственно распределенную по функциям и средствам иерархическую структуру. АСУ ТП включает в себя следующие уровни иерархии (рис. 2):

1. Уровень связи с ТООУ (датчики, пусковые устройства) — обеспечивает подготовку и проведение автоматических измерений параметров технологического процесса, контроля состояния оборудования, отработку управляющих команд автоматического и автоматизированного управления и защит (подача силового питания на исполнительные механизмы).
2. Нижний уровень автоматизации — обеспечивает прием и обработку данных автоматических измерений и контроля, обмен информацией с верхним уровнем, АСУ ТП, осуществляет реализацию алгоритмов управления системами безопасности, технологических защит и блокировок, автоматического и автоматизированного (по командам с БПУ и РПУ)
3. Верхний уровень АСУ ТП — обеспечивает обработку информации, ее архивирование, документирование и представление оперативному персоналу БПУ и РПУ, осуществляет формирование команд автоматизированного управления технологическим процессом энергоблока, обеспечивает передачу необходимой информации от АСУ ТП ЭБ через УПД (см. 2) в сеть АСУ П и СПНИ, а через коммутаторы межсерверного обмена в систему верхнего общестанционного уровня (СВСУ) для АРМ начальника смены станции (НСС) и АРМ локального кризисного центра, а также от СРВПЭ в АРМ СРВПЭ, размещенном в помещении ЛКЦ защищенного пункта управления (ЗПУ).

В дополнении к этому в проекте реализованы следующие основные критерии проектирования:

- сведение к минимуму вероятности снижения мощности из-за единичной неисправности в АСУ ТП;
- принятие мер по снижению вероятности ошибок персонала при подаче команд управления;
- подавление неактуальных сигналов.

2.8. Функция СППБ

Назначение функции состоит в поддержке операторов по принятию решений и выполнению противоаварийных действий.

СППБ предназначена для обеспечения выполнения оперативным персоналом действий по предотвращению и ограничению последствий аварий, включая запроектные.

СППБ должна обеспечивать снижения нагрузки на персонал и вероятности его ошибок при выполнении задач управления при авариях.

СППБ предназначена для быстрого отображения необходимой информации при нарушении КФБ и оказания поддержки принятия решений оперативным персоналом при выполнении процедур восстановления КФБ с учетом «Руководства по управлению запроектными авариями» (РУЗА).

Задачи СППБ:

- Своевременное уведомление оператора о нарушении КФБ, отображение параметров безопасности, характеризующих состояние КФБ. Контроль и регистрация изменений состояния КФБ).
- Представление оперативному персоналу электронного аналога деревьев состояния КФБ, определяющих выбор процедуры по восстановлению КФБ.
- Выдача оператору рекомендаций о применении той или иной процедуры по восстановлению КФБ (Контроль состояния диагностических блоков (узлов) деревьев КФБ. Контроль активации выходных ветвей деревьев КФБ).
- Предоставление электронных интерактивных процедур восстановления КФБ.
- Предоставление, справочной или расчетной информации, необходимой для принятия решений и выполнения управляющих действий, определенных в «Руководстве по управлению запроектными авариями».
- Предоставление необходимой информации о достижении целей, на которые направлены действия персонала.

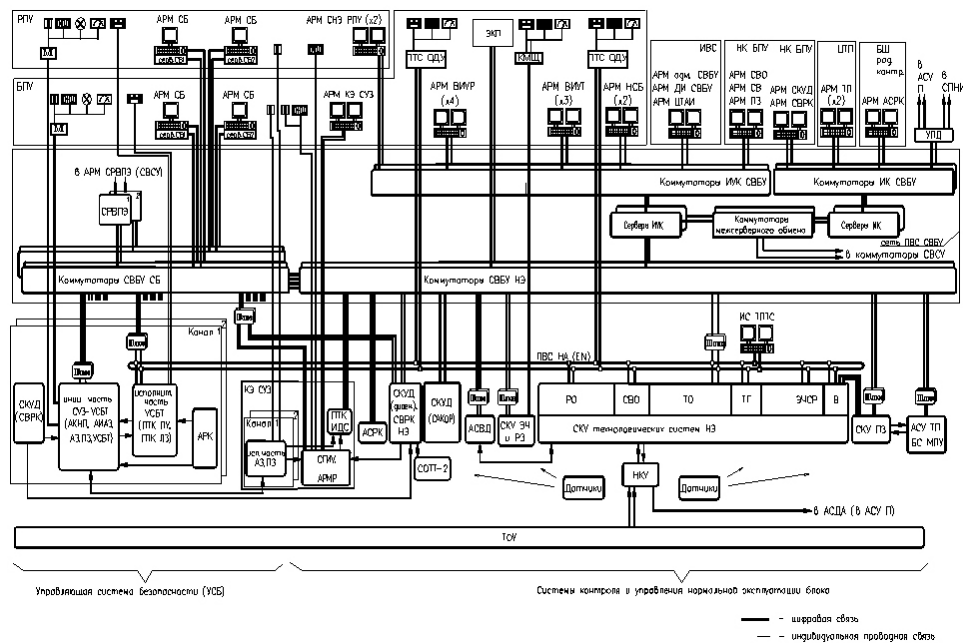


Рис. 2: Упрощенная структурная схема АСУ ТП

Структура СППБ:

- СППБ является частью системы информационной поддержки оператора, которая в свою очередь является частью верхнего уровня АСУ ТП блока.
- СППБ реализуется отдельной трехуровневой подсистемой:
 - 1) 1-ый уровень отображает статус каждой из КФБ и параметры безопасности;
 - 2) 2-ой уровень — деревья состояний КФБ, а также фрагменты, представляющие оператору информацию для принятия решений по восстановлению КФБ;
 - 3) 3-ий уровень — поддерживающие форматы, обеспечивающие оператора дополнительной информацией, необходимой для выполнения конкретных управляющих действий или для выяснения причин невозможности их выполнения.
- СППБ должна сохранять работоспособность во всех предусмотренных рабочим технологическим регламентом режимах.

АСУ ТП состоит из ряда подсистем выделенных по технологическому или функци-

ональному признаку. Все системы объединяются в единую систему с помощью системы верхнего блочного уровня (СВБУ), которая обеспечивает сбор и обмен информации, дистанционное управление предусмотренного проектом состава оборудования, а также реализовывает общешлюзовые задачи.

В составе систем автоматизации НЭ предусматривается обобщенная мнемосхема, включающая индивидуальные приборы, индикаторы положения основного оборудования и табло, и средства индивидуального управления для возможности быстрой оценки состояния энергоблока, а также возможности его эксплуатации на мощности в течение ограниченного времени и планового останова при ситуации связанной с потерей СВБУ.

Проект предусматривает централизацию контроля и управления энергоблоком.

В состав пунктов управления и контроля энергоблока входят:

- централизованные пункты;
- блочный пункт управления (БПУ);
- резервный пункт управления (РПУ);
- блочный пункт радиационного контроля (БПРК);

— местные пункты управления (контроля).

Оперативный контроль и управление блоком осуществляется с БПУ, на котором предусмотрено постоянное пребывание трех операторов: НСБ, ВИУР и ВИУТ. При необходимости может быть привлечен дополнительный персонал к зоне вспомогательных систем и электрической части. На БПУ энергоблока для контроля и управления энергоблоком предусмотрены необходимые средства автоматизации, включая автоматизированные рабочие станции (АРМ), средства индивидуального контроля и управления и экран коллективного пользования (ЭКП).

Основной режим работы оператора — контроль и управление с мониторов рабочих станций. При этом осуществляется контроль и управление, как системами нормальной эксплуатации, так и системами безопасности. При работе защит безопасности накладывается запрет на управление с мониторов рабочих станций оборудованием систем безопасности.

На операторов возлагаются задачи ввода в работу и вывода из работы систем и отдельного оборудования (при изменении режима блока), контроля за работой блока, задания режимов работы оборудования, регуляторов, АВР, ФГУ и периодического опробования СБ, а также автоматизированного дистанционного контроля и управления оборудованием при аварийных режимах работы (при необходимости такого управления) и в условиях, когда может потребоваться автоматизированное дистанционное управление.

Операции, требующие немедленного выполнения при создании соответствующих условий, реализуются автоматически с выдачей оперативному персоналу необходимых сообщений.

2.9. Оценка текущей ситуации на блоке

Для быстрой оценки всеми операторами текущей ситуации на блоке, БПУ осна-

щен экраном коллективного пользования (ЭКП) с выводом на него технологических видеокadres текущего режима с составом информации, которая должна однозначно информировать персонал о режиме работы блока и имеющихся место нарушениях. Информация на ЭКП выводится от СКУ НЭ минуя серверы СВБУ. На рис. 3 представлена компоновка пульта БПУ. На рис. 4 представлен общий вид пульта БПУ.

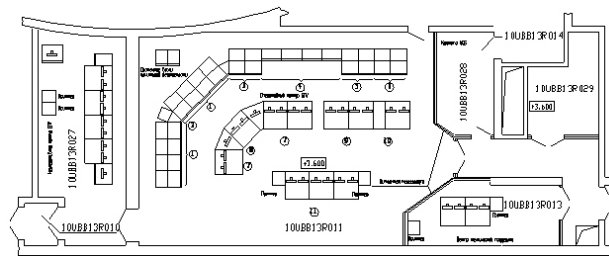


Рис. 3: Компоновка пульта БПУ



Рис. 4: Общий вид пульта БПУ

АСУ ТП создается как единая цифровая система контроля и управления, состоящая из УСБ и УСНЭ. При этом АСУ ТП реализуется на современных программно-технических средствах.

Основными целями создания СВБУ являются обеспечение:

- контроля и управления технологическим процессом;
- интеграция всей информации по энергоблоку от всех систем и подсистем АСУ ТП;
- дисплейного управления оборудованием систем нормальной эксплуатации и оборудованием систем безопасности в режимах, предусмотренных проектом;

- централизованного контроля и представления, как обобщенной, так и детализированной информации о состоянии энергоблока, отдельных параметрах технологического процесса и состоянии оборудования (арматуры и механизмов);
- контроль состояния барьеров безопасности энергоблока;
- необходимой информацией персонала различных подразделений АЭС, которым эта информация необходима в процессе работы;
- формирования сигнализации о нарушениях в работе энергоблока, отдельных систем, отдельного оборудования;
- необходимой информацией Аварийного центра АЭС;
- возможности обмена информацией с СУЗ, УСБТ, СКУД (СВРК, САКОР, СКД), СКУ НЭ, СВРК, АСВД, САКОР, АКНП, СКУ ВХР, АСРК, СКУ ПЗ, СКУ СВО, СКУ ЭЧ, СРВПЭ, СВСУ, СПНИ;
- ведения архивов;
- возможности обмена информацией СВБУ с общестанционным уровнем;
- единого подхода в организации информационного обеспечения.

Основными направлениями достижения указанных выше целей является использование:

- современного программного обеспечения на базе программной платформы «ПОРТАЛ», обеспечивающего возможность обработки оперативных данных, получения на этой основе информации о состоянии технологических процессов и оборудования, характере и причинах нарушений и позволяющего операторам ЭБ принимать адекватные решения по управлению оборудованием ЭБ;
- современных компьютерных технологий, обеспечивающих эффективный человеко-машинный интерфейс;
- интеграция информации от всех подсистем АСУТП в едином информационном пространстве;

- структурирование информации и развитие информационной поддержки оперативного персонала.

Границы системы проходят по выходным разъемам шлюзов подсистем, информация от которых интегрируется в СВБУ. Шлюзы со смежными системами в состав СВБУ не входят.

В СВБУ должны быть реализованы информационные, управляющие и вспомогательные функции.

К информационным функциям СВБУ относятся:

- контроль текущего состояния ТОО;
- оперативный анализ состояния ТОО;
- представление информации;
- регистрация и архивирование.

К управляющим функциям относятся:

- дистанционное управление оборудованием с АРМ СВБУ;
- автоматизированное управление в части задания режимов работы оборудования НЭ, ввода/вывода защит, перевода регуляторов в режим дистанционного управления или авторегулирования, запуска пошаговых программ пуска/останова групп оборудования (групповое, подгрупповое управление).

К вспомогательным функциям относятся:

- оперативный анализ состояния оборудования АСУ ТП по отдельным системам с глубиной, определяемой на стадии «Технического проекта»;
- программное и операторское управление работой СВБУ;
- самодиагностика;
- контроль и защита от несанкционированного доступа.

Минимальные характеристики системы должны обеспечивать прием и обработку следующего количества параметров (табл. 2).

В состав АСУ ТП АЭС входят более 20 подсистем, разработанные различными производителями, том числе рядом предприятий Росатома таких как:

Таблица 2: Требования к техническим характеристикам по приему информации

Вид технологического объекта	Всего не менее
Аналоговые технологические переменные	10 000
Дискретные технологические переменные	20 000
Расчетные параметры, выполняемые в рамках СВБУ (данный параметр должен быть уточнен на стадии рабочего проектирования СВБУ)	10 000
Диагностические переменные	100 000

ФГУП «ВНИИА им. Н.Л. Духова», ФГУП «ФНПЦ НИИС» им. Ю.Е. Седакова», ФГУП «РФЯЦ-ВНИИЭФ», ФГУП ФНПЦ «ПО «Старт» им. М.В. Проценко», УЭМЗ, ФГУП «ПСЗ», ФГУП «РФЯЦ-ВНИИТФ им. академ. Е.И. Забабахина».

3. Выбор технологий решения задач

Для реализации функций обучения и информационно-справочной поддержки пользователей АСУ ТП АЭС на современном этапе наиболее перспективным решением является создание программно-технического комплекса (ПТК) для типового выставочного стенда (ВС) АСУ ТП АЭС и системы класса 4i — Интегрированная Интерактивная Интеллектуальная Информационно-справочная, обучающая и выставочная система (ИАС4i) [3].

ИАС4i — это информационно-справочная, обучающая и демонстрационная система, имеющая развитые возможности представления экспонатов на основе мультимедийного представления информации, применения виртуальной реальности.

Типовая демонстрационная система обеспечивает «погружение» участников выставочного мероприятия в информационную среду демонстрируемой проблематики с учётом интересов (подготовленности аудитории), как для коллективного просмотра, так и по индивидуальным запросам пользователей [2].

4. Направления исследований

РФФИ поддержан в 2017 году грант на 2017-2019 гг за № 17-07-00833 «Разработка архитектуры и прототипа типовой ИАС для

демонстрации возможностей и обучения решениям АСУТП АЭС на основе системы виртуальной реальности».

В рамках указанного гранта предстоит провести исследования по созданию прототипа автоматизированной системы, предназначенной для решения задач информационной поддержки демонстрации возможностей и внедрения АСУ ТП АЭС.

Литература

- [1] Менгазетдинов Н.Э., Полетыкин А.Г., Промыслов В.Г., Зуенкова И.Н., Бывайков М.Е., Прокофьев В.Н., Коган И.Р., Коршунов А.С., Фельдман М.Е., Кольцов В.А. Комплекс работ по созданию первой управляющей системы верхнего блочного уровня АСУТП для АЭС «Бушер» на основе отечественных информационных технологий. М.: ИПУ РАН, 2013. 95 с.
http://www.ipu.ru/sites/default/files/page_file/busher.pdf
- [2] Захарушкин В.Ф. Актуальные вопросы этапа запуска проекта при выполнении НИОКР. Труды Международной научной конференции «Ситуационные центры и ИАС4i для мониторинга и безопасности» SCVRT2013-14, 25-28 ноября 2013-2014, Протвино, Московская область, Россия Изд.ИФТИ, Протвино-Москва, 2014 2013-2014, ISBN 978-5-88835-027-0, с. 1-3.
- [3] Алёшин В.П., Астахов Ю.С., Борисов Т.Н., Захарушкин В.Ф. и др. Пилотный проект центрального ситуационно-аналитического центра для мониторинга и управления объектами железнодорожной инфраструктуры. Труды второй научно-технической конференции «Интеллектуальные системы управления на железнодорожном транспорте (ИСУЖТ-2013)», 21-22 октября 2013г, Москва, Россия (ОАО «НИИАС»), с. 116-123.

Выбор аппаратно-технических средств для реализации системы дополненной реальности для задач АСУ ТП*

*В. Ф. Захарушкин*¹, *М. А. Хламов*²
 zakharushkin@mail.ru, mkhlamov@gmail.com

¹Протвино, Институт физико-технической информатики

²Московский физико-технический институт (государственный университет)

В статье рассматриваются современные аппаратные и программные средства для создания систем дополненной реальности. Актуальность задачи обусловлена необходимостью использования преимуществ дополненной реальности в задачах АСУ ТП АЭС. Рассматриваются существующие системы AR и дается их сравнительная характеристика в контексте рассматриваемой задачи.

Ключевые слова: дополненная реальность, АСУ ТП, AR.

Hardware and software tools for augmented reality system for Automated Control Systems*

*V. F. Zakharushkin*¹, *M. A. Khlamov*²

¹Protvino, Institute of Computing for Physics and Technology

²Moscow Institute of Physics and Technology (State University)

This paper considers modern hardware and software tools used to create augmented reality systems. The urgency of the problem is caused by the need of using augmented reality systems advantages in the ACS TP of nuclear power station. We consider the existing AR systems and give the comparative analysis in the context of the stated problem.

Keywords: augmented reality, automated control systems, AR.

Введение

В последнее время активно развиваются технологии дополненной и виртуальной реальности. Дополненная реальность — это целая методика, позволяющая дополнять реальный мир новой информацией с помощью цифровых технологий. Ученый Рональд Азума определил в 1997 году, что технология дополненной реальности — это система, совмещающая виртуальное и настоящее, взаимодействующая в текущем времени и работающая в 3D [1]. Дополненная реальность в настоящее время реализуется двумя способами: с использованием специального шлема (очков) (см. рис. 1), содержащего в себе или связанного с компьютером, или с помощью имеющегося гаджета.

Начиная с 2008 года AR становится одним из основных трендов информационных технологий [2], что связано с широким распространением мобильных компьютерных решений, оснащённых качественной видео-



Рис. 1: Шлем дополненной реальности

камерой и датчиками позиционирования. Быстро растущая производительность процессоров для мобильных приложений делает возможным рендеринг изображений виртуальных объектов (включая трёхмерные) и их наложение на картинку видеокамеры с учётом её положения в пространстве, что в сочетании с доступом в Интернет открывает широкие возможности для использования технологии AR в различных сферах [3]. Существует достаточно широкий спектр областей науки и техники, в которых может

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 17-07-00833, 18-07-20030.

применяться дополненная реальность [3], однако в первую очередь можно выделить следующие из них:

- медицина;
- проектирование и дизайн;
- картография и геоинформационные системы (ГИС);
- реклама;
- образование;
- игровая индустрия;
- военная техника.

Как видно из перечня, по сути, любая предметная область, где есть необходимость смотреть, может стать объектом применения AR-технологии.

Аппаратные платформы дополненной реальности

На рис. 2 представлена схема развития платформ дополненной реальности [4].



Рис. 2: Схема развития AR платформ

По типу используемых платформ все AR приложения можно разделить на два типа:

- AR с использованием смартфона или планшета;
- AR с использованием специальных устройств, в основном — шлемов дополненной реальности.

Приложения первого типа доступны широкой аудитории. У обоих гигантов на рынке ОС для смартфонов есть свои платформы для разработчиков: ARKit у Apple и

ARCore у Google. Технические характеристики таких систем ограничены возможностями используемого смартфона или планшета. Основные недостатки таких систем:

- a) необходимость все время держать перед собой устройство и таким образом ограничить использование рук;
- b) низкая точность работы датчиков на некоторых устройствах;
- c) маленький угол обзора.

AR на смартфонах и планшетах активно используются различными компаниями для визуализации своих товаров. Ярким примером является приложение от IKEA [5]. Также такие приложения есть и в ресторанной индустрии [6], косметике [7], автомобильной промышленности [8] и других.

Эта технология мало подходит для длительного использования, например в обучении. Для этого перспективным направлением в AR более подходящим видится использование специального оборудования. Существует три типа AR-шлемов: автономные, требующие подключения к ПК, требующие подключения к смартфону. Основной их недостаток — цена, которая может достигать до нескольких тысяч долларов за штуку.

Microsoft HoloLens

Microsoft HoloLens [9] — первый шлем смешанной реальности (Mixed Reality) и первый продукт Microsoft, работающий в операционной системе Windows 10 и подсистеме Windows Holographic (см. рис. 3).



Рис. 3: Microsoft HoloLens

HoloLens представляет собой надеваемый на голову обруч с расположенными перед

Таблица 1: Очки дополненной реальности

Производитель	Microsoft	Meta	ODG	Sony
Модель	Microsoft HoloLens	Meta 2	R-7 Smartglasses System	SmartEyeglass SED-E1
Процессор	Intel Atom X5-Z8100	Intel® Core™ i7-6700 Processor or AMD AMD FX™ 9590, equivalent or better	Qualcomm® Snapdragon™ 805 2.7GHz quad-core Processor	Нужно соединение со смартфоном
Оперативная память	2 ГБ	16 GB RAM DDR4 or more	3GB	Нужно соединение со смартфоном
Аккумулятор	16500 мВч	-	Dual 650mAh Batteries	150 min
Среда разработки	Unity (Vuforia SDK), DirectX, Windows API	Unity 5.6 or higher	Android API, Unity (Vuforia), Augumenta, EON Reality, Wikitude	Android API
ОС	32-рядная Windows 10.0.11802.1033	Windows 10 (64-bit)	ReticleOS (based on Android 6)	Android 4.4
Дополнительно	Встроенные микрофоны и камеры, возможность голосового управления	Массив сенсоров для трекинга рук и отслеживания положения в пространстве, включая шестиосный IMU	Акселерометр, гироскоп, магнетометр	Акселерометр, гироскоп, компас, микрофон
Поле обзора	30°	90°	30°	23°
Стоимость	Development Edition: 3000\$ Commercial: 5000\$	1495\$	2750\$	2000\$
Разрешение дисплея	1268 × 720 (один глаз)	2560 × 1440	1280 × 720	419×138 monochrome
Требуется подключение к внешнему устройству	нет	да	нет	да
Камера	1 megapixel	1 megapixel	2 megapixel	3 megapixel

глазами тонированными линзами с волнообразной призматической структурой, которые преломляют и отправляют в глаза пользователя изображения с расположенных по бокам микродисплеев. Целевой аудиторией проекта является бизнес, и решениями с использованием Hololens пользуются такие организации, как Japan Airlines, NASA, Volvo, Autodesk, Trimble и другие.

Meta 2

Meta 2 Development Kit [10] — это второе поколение очков Meta, предназначенных для разработчиков программного обеспечения, и представитель нового типа компьютерных интерфейсов. Они подключаются к компьютеру вместо монитора, визуализируют двухмерные и трёхмерные данные подобно голограммам, которые можно привязать к определённому месту в окружающем пространстве, а управление осуществляется с помощью привычных органов управления (мыши и клавиатуры), жёсткими руками, через движения пользователя (в очки встроены датчики положения) и го-

лосом. Новинка отличается усовершенствованиями, коснувшимися как дизайна, так и функционала — в том числе, расширенным полем обзора, улучшением считывания жестов, позиционного трекинга и SDK (см. рис. 4).



Рис. 4: Meta 2 Development Kit

Magic Leap

Magic Leap [11] состоит из трех частей: самих AR-очков (Lightwear), вычислитель-

Таблица 2: Очки дополненной реальности (продолжение)

Производитель	Vuzix	Vuzix	Epson Moverio	Epson Moverio
Модель	Vuzix M300 Smart Glasses	Vuzix M100 Smart Glasses	BT-300	BT-350
Процессор	Dual Core Intel Atom CPU	OMAP4460 at 1.2GHz	Intel® Atom™ x5, 4 cores	
Оперативная память	2 ГБ	1GB	2GB	2GB
Аккумулятор	Внутренняя батарея - 160mAh, внешняя - 860mAh, 2 - 12 часов работы	3800 mAh	2950 mAh	
Среда разработки	Unity (Vuforia)	Unity (Vuforia)	Android API, Unity (Official Epson plugin)	
ОС	Android 6 OS	Android ICS 4.04, API 15	Android 5.1	
Дополнительно	Распознавание голоса, трекинг головы, гироскоп, акселерометр, компас, GPS	Распознавание жестов, GPS, распознавание голоса, трекинг головы, гироскоп, акселерометр, компас	GPS, компас, гироскоп, акселерометр, микрофон	
Поле обзора	16.7°	15°	23°	
Стоимость	1700 \$	1079\$	700\$	2000\$
Разрешение дисплея	640 × 360	400 × 240	1280 × 720	
Требуется подключение к внешнему устройству	нет	нет	нет	нет
Камера	10 megapixel	5 megapixel	5 megapixel	5 megapixel

ного блока-шайбы со встроенной батареей (Lightpack), закрепляемого на поясе, и беспроводного контроллера (Control). Аппарат начнут поставлять в 2018 году (см. рис. 5).



Рис. 5: Magic Leap

Спецификации Magic Leap One держатся в секрете. Судя по иллюстрациям, Lightwear будут подключаться к Lightpack по проводу, однако внешний вид и комплектация коммерческого продукта могут быть изменены. Как говорит глава компании Рони Абовиц, гаджет будет комплектоваться отдельной графической картой, четырьмя встроенными микрофонами, как минимум шестью внешними камерами, а также встроенными динамиками, подобно HoloLens. Производительность устройства будет сопоставима с MacBook Pro или игровым компьютером Alienware.

Magic Leap — один из самых многообещающих, но скрытных стартапов. В перспективность его разработок верят крупнейшие игроки IT-рынка: в фирму вложили почти \$2 миллиарда такие инвесторы, как Google, Alibaba, KPCB и Andreessen Horowitz.

Epson Moverio

Epson [12] предоставляет 4 актуальные модели очков для дополненной реальности: Epson Moverio BT-300 (см. рис. 6), Epson Moverio BT-350, Epson Moverio Pro BT-2000, Epson Moverio Pro BT-2200.



Рис. 6: Epson Moverio BT-300

Еще осенью 2016 года компания Epson начала сотрудничество с DJI. Плодом же их совместной работы стало приложение DJI GO, которое идеально подходит для управления дроном. Также очки от Epson можно использовать для просмотра развлекательного контента, обучения, в бизнесе и промышленном производстве.

Vuzix M300 Smart Glasses

Vuzix Corporation была основана в 1997 году и является разработчиком и поставщиком смарт-очков и видеогарнитур на потребительский и промышленный рынки.



Рис. 7: Vuzix M300 Smart Glasses

M300 Vuzix основаны на предыдущей модели M100 с улучшенными характеристиками. Особенностью является использование сенсорной панели на боковой стороне вместо смартфона для управления устройством. Также можно использовать жесты или голосовое управление.

Направления создания AR в АСУ ТП АЭС

Перспективными направлениями развития технологии VR/AR в АСУ ТП являются создание систем:

- демонстрационной и выставочной систем;
- аналитической и справочной систем проектирования;
- обучающих систем дополненной реальности персонала по обслуживанию и ремонту технических средств.

Анализ перспективных аппаратно-программных платформ для создания систем AR в АО «РАСУ»

Технические характеристики аппаратных платформ приведены в сводной таблицах 1 и 2.

В таблицах 3 и 4 приведена информация о программных платформах, используемых для разработки дополненной реальности.

Стоимость приложения зависит от количество предоставляемого функционала. У большинства продуктов существует так называемая Pro-версия с максимальным предоставляемым функционалом, стоимость которой необходимо обговаривать отдельно.

В то время как для демонстрационных и выставочных систем возможно использование решений на смартфонах или планшетах, то для других важна свобода действий пользователя. Поэтому далее мы рассмотрим общие требования к аналитическим, справочным и обучающим системам.

В 2017 году организации UI LABS [13] и AREA (The Augmented Reality for Enterprise Alliance) [14] объявили о публикации первого в мире руководства, в котором сформулированы функциональные требования к аппаратной и программной частям решений в области дополненной реальности, которые позволят создавать такие решения, адресованные промышленным потребителям [15]. Всего в разработке документа приняли участие 65 предприятий и организаций, например, компании Microsoft, General Electric, Rolls-Royce, Dow Chemical, Intel, the US Air Force, Boeing и другие. Разработанные рекомендации охватывают время автономной работы, средства подключения, поле зрения, встроенную память, встроенную ОС, средства ввода-вывода, физические характеристики и требования безопасности. Также затронуты вопросы формирования и развертывания контента, связи с интернетом вещей. Эти рекомендации призваны служить ориентиром для производителей аппаратного и программного обеспечения. На данный момент не существует оборудования, полностью соответствующего требованиям, указанным в документе. Однако на его основе можно сформулировать требования к разрабатываемым системам.

В таблице 5 приведены требования к аппаратной части системы дополненной реальности для АСУ ТП.

Таблица 3: Программные платформы для разработки AR

	Vuforia	EasyAR	Wikitude
Платформы	Android, iOS, UWP, Unity Editor	Android, iOS, UWP, Windows, Mac, Unity Editor	Android, iOS, Smart Glasses. Unity
Умеет распознавать	3D объекты (куб, цилиндр, плоскость), текст, VuMark (комбинация картинки и QR-кода), изображения	QR-коды, изображения, 3D объекты (в платной версии)	3D объекты, изображения, SLAM, GPS
Минимальный угол, требуемый для распознавания	30	35	40
Максимальное расстояние, на котором распознается маркер (м)	1.2	0.9	0.8
Поддерживаемые API	Unity, C++, Java	Unity, C/C++, Java, Swift, Objective-C	Unity, Cordova/PhoneGap, Titanium, Xamarin
Стоимость	Бесплатно – 499 \$ за приложение – 99 \$ в месяц	Бесплатно – 499 \$	Бесплатно – 2490 € – 2990 € – 4490 €

Таблица 4: Программные платформы для разработки AR (продолжение)

	ARToolKit	Kudan	Maxst
Платформы	Android, iOS, Linux, Windows, Mac OS and Smart Glasses.	Android, iOS	Android, iOS, Windows, Mac OS
Умеет распознавать	Черные квадраты, плоские изображения	Изображения, 3D объекты, SLAM	Изображения, 3D объекты, QR/Barcode сканнер, SLAM
Минимальный угол, требуемый для распознавания	10	30	50
Максимальное расстояние, на котором распознается маркер (м)	3	0.8	0.5
Поддерживаемые API	Unity, OpenSceneGraph	Unity	Unity, Android, iOS
Стоимость	Бесплатно – open source	Бесплатно – 1000 \$ в год	Бесплатно – 499 \$ – 599 \$

Таблица 5: Требования к аппаратной части

Аккумулятор	8 часов непрерывной работы (4 при возможности заменить батареею)
Подключения	Bluetooth, Wi-Fi 802.11
Угол обзора	не менее 20°
Хранилище	8Гб
Процессор	2-ядерный, 1,2 ГГц или лучше
Оперативная память	2 Гб
Разрешение дисплея	1280 × 720
Камера	2 мегапикселя
Дополнительно	GPS, компас, гироскоп, акселерометр, микрофон

Таким образом, для реализации системы дополненной реальности для АСУ ТП подходят следующие AR-шлемы:

- Epson Moverio BT-350;
- ODG R-7 Smartglasses System;
- Vuzix M300 Smart Glasses.

Для работы приложения дополненной реальности необходимо обнаружить реальный объект в видеопотоке и затем «до-

полнить» его. Существуют различные виды таких реальных объектов, которые можно пытаться обнаружить на видеопотоке с камеры.

А. Маркер.

В работе [16] приведена краткая схема обработки изображения в AR-приложениях, использующих в качестве маркера черно-белое изображение (см. рис. 8).

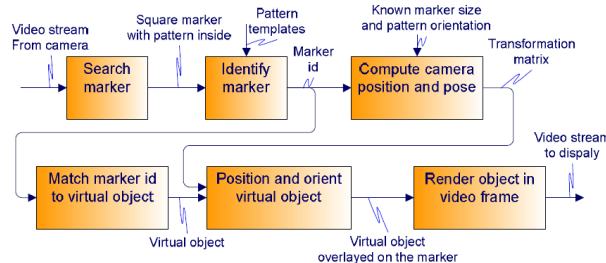


Рис. 8: Схема обработки изображения в AR-приложении

Это самый простой и надежный способ реализации дополненной реальности, однако он требует специальной предварительной подготовки (нужно распечатать, вырезать и разместить эти маркеры).

- В.** Безмаркерная дополненная реальность. На самом деле, при таком подходе все равно используется маркер, в качестве которого выступает картинка. Для успешного использования этого способа необходимо иметь достаточно детализированное цветное изображение с соотношением сторон близким к 1:1. Затем можно реальные 2D объекты использовать в качестве маркера незаметно для пользователя, например «оживлять» иллюстрации в книге.
- С.** 3D объекты. Также есть возможность использовать в качестве маркеров простые объемные объекты, такие как куб и цилиндр.
- Д.** GPS. Некоторые приложения не используют данные видеопотока, а используют данные, предоставляемые GPS-приемником, компасом, акселерометром. На их основе рассчитывается положение и поворот пользователя в пространстве.
- Е.** SLAM. SLAM — Simultaneous Localization and Mapping (Метод одновременной локализации и построения карты) — метод, используемый для построения карты в неизвестном пространстве или для обновления карты в заранее известном пространстве с одновременным контролем текущего местоположения и пройденного пути. На данный момент приложения на мобильных устройствах умеют определять вертикальные и горизонтальные поверхности и прикреплять к ним объекты. Перспективны данной технологии в сочетании с распознаванием образов и машинным обучением огромны, однако упираются в ограничения используемых устройств.

В таблице 6 приведены требования к программной части системы дополненной реальности для АСУ ТП с учетом выбранных аппаратных платформ.

Таблица 6: Требования к программной части

Операционная система	Android
Платформы	Android API, Unity
Маркеры	Монохромные и цветные изображения

Среди рассмотренных программных платформ лучшими являются:

- Vuforia;
- Wikitude.

Выводы

По результатам проведенного исследования авторы пришли к следующим выводам:

1. Технологии VR/AR являются быстро развивающимся направлением применения информационных технологий в различных сферах деятельности.
2. Системы АСУ ТП АЭС являются одной из сфер деятельности, которые могут эффективно применять системы VR/AR.
3. Наиболее перспективными направлениями развития технологии дополненной реальности являются аналитические, справочные и обучающие системы.
4. Наиболее перспективными AR-платформами являются Epson Moverio BT-350 и ODG R-7 Smartglasses System в сочетании с Vuforia.
5. Следующим этапом работ является выбор бизнес-процессов для реализации пилотного проекта.

Литература

- [1] Ronald T. Azuma (1997). A survey of augmented reality. Presence: Teleoperators and Virtual Environments, Volume 6 Issue 4, August 1997, стр. 355-385.
- [2] Сергей Солдатов, Нина Кузьмина (2016). Интерфейс будущего — системы дополненной реальности. Современные технологии автоматизации, стр. 96-103.
- [3] Авдошин А.С., Забержинский Б.Э., Головин К.Ю. (2012). Анализ возможностей и перспектив

- использования дополненной реальности в теории и на практике. Актуальные проблемы науки, экономики и образования XXI века: матер. II Международной научно-практ. конф., 5 марта — 26 сент. 2012 г. — Самара: Самарский ин-т (фил.) РГТЭУ, 2012.
- [4] Блог компании friifond. Рынок виртуальной и дополненной реальности: перспективы для стартапов с точки зрения инвестора, 2017 Web: <https://habrahabr.ru/company/friifond/blog/322230>
- [5] IKEA Place app official website, Web: <https://www.ikea.com/gb/en/customer-service/ikea-apps>
- [6] Kabaq official website, Web: <http://www.kabaq.io/>
- [7] MAKEUP GENIUS official website, Web: <https://campaign.loreal-paris.ru/makeupgenius/>
- [8] Hyundaiar download website, Web: <https://play.google.com/store/apps/details?id=com.rrd.hyundaiar>
- [9] Microsoft Hololens official website, Web: <https://www.microsoft.com/en-us/hololens>
- [10] Metavision official website, Web: <http://www.metavision.com/>
- [11] Magicleap official website, Web: <https://www.magicleap.com/>
- [12] Epson Moverio official website, Web: <http://www.epson.ru/catalog/moverio/>
- [13] Ullabs official website, Web: <http://www.uilabs.org/>
- [14] The AREA official website, Web: <http://thearea.org/>
- [15] Augmented Reality Functional Requirements, Web: <http://thearea.org/area-resources/augmented-reality-functional-requirements/>
- [16] Juha Sääsäski, Tapio Salonen, Marja Liinasuo, Jarkko Pakkanen, Mikko Vanhatalo, Asko Riitahuhta (2008). Augmented reality efficiency in manufacturing industry: a case study. Proceedings of NordDesign 2008, August 21-23, 2008, Tallinn University of Technology, Tallinn, Estonia.

Уровневое моделирование на основе комплексного подхода*

Д. А. Завьялов, А. А. Захарова

zda@tpu.ru | zaa@tpu.ru

Национальный исследовательский Томский политехнический университет

Обоснована актуальность применения комплексного подхода и уровневого моделирования при проектировании сложных систем на примере месторождений нефти и газа. Предложен комплексный подход, заключающийся в использовании разноуровневых и разномерных моделей. Показана эффективность такого подхода.

Ключевые слова: комплексный подход, уровневое моделирование, принятие решений, геологическая модель, гидродинамическая модель, проектирование разработки месторождений

Level modeling based on an integrated approach*

D. A. Zavyalov, A. A. Zakharova

¹National Research Tomsk Polytechnic University

The urgency of applying the integrated approach and level modeling in the design of complex systems based on the example of oil and gas fields is substantiated. A comprehensive approach is proposed, consisting of using different-level and different-sized models. The effectiveness of this approach is shown.

Keywords: complex approach, level modeling, decision making, geological model, hydrodynamic model, field development design.

Введение

Сложные системы характеризуются большими объемами данных, накопление которых происходит на протяжении всего жизненного цикла. При этом таким системам свойственна неоднородность данных, описывающих отдельные их составляющие. Моделирование и принятие решений требуют учета всех данных, их верификации и оценки достоверности моделей при моделировании на разных уровнях детализации.

Комплексный подход при моделировании сложных систем

Комплексный подход при моделировании сложных систем заключается в использовании моделей разных типов и уровней в зависимости от достоверности и объемов данных, выполнении верификации исходных данных и оценки достоверности моделей для контроля ошибок на всех стадиях проектирования.

Высокой точности моделирования невозможно добиться на ранних стадиях жиз-

ненного цикла системы, в тоже время высокая изученность на более поздних стадиях позволяет строить достаточно достоверные и подробные модели. Однако большинство проектировщиков, имея в своем распоряжении достаточные вычислительные мощности, создают модели максимальной размерности, даже когда для этого недостаточно данных. Для повышения эффективности процесса проектирования необходимо снижать детализацию моделей, сохраняя при этом точность результата моделирования, то есть необходим баланс между детализацией модели и точностью моделирования.

Поэтому необходимо использовать комплексный подход в моделировании, который предполагает:

- снижение избыточности,
- снижение ресурсоемкости,
- снижение временных затрат,
- не повлияет на результат моделирования (в рамках допустимой погрешности, которая определяется поставленной задачей).

Предлагается объединить все модели в единую комплексную модель (рис. 1), уровень и детализация которой определится решаемой задачей, объемом имеющихся дан-

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 18-07-20030, 18-37-10008 и Министерства образования и науки РФ, грант 2.1642.2017/4.6.

ных [1]. Кроме того, в ряде задач необходимо по возможности переходить от численных моделей к визуальным [2, 7, 8].

Жизненный цикл и роль цифровых моделей в проектировании разработки месторождений

Примером сложной системы является месторождение углеводородов [3, 4], состоящее из залежей нефти, скважин, наземной инфраструктуры. Процесс проектирования его разработки порождает противоречие между целями и интересами проектировщика, компании-недропользователя и государства. Поэтому эффективность разработки месторождения зависит от многих факторов и определяется не только экономическими показателями или полнотой извлечения запасов. На протяжении своего жизненного цикла месторождение проходит 5 основных стадий: открытие, оценка запасов, подготовка к освоению, добыча, ликвидация. Изученность месторождения различается и определяется стадией жизненного цикла. Несмотря на жесткие ограничения, установленные действующими регламентами [5, 6], необходимо применять различные алгоритмы и методики при проектировании разработки в зависимости от изученности месторождения.



Рис. 2: Зависимость погрешности модели от количества информации

В течение всего жизненного цикла месторождения происходит постоянное накопление данных о его структуре и строении, ис-

тории разработки, поэтому изученность месторождения напрямую зависит от стадии жизненного цикла. В свою очередь изученность месторождения определяет точность и достоверность его моделей (рис. 2), которые создаются для составления прогноза разработки.

Для добычи углеводородов недропользователь должен иметь действующий проектный документ на разработку месторождения, утвержденный государственной комиссией. Такой проектный документ устанавливает объемы добычи, которые должны быть соблюдены недропользователем. При этом ядром любого проекта на разработку является трехмерная модель месторождения. Такая модель (рис. 3) представляет собой цифровое отображение ряда свойств реального пласта.

В процессе создания проекта на разработку принимают участие различные специалисты: сейсмик, геофизик, геолог, специалист по моделированию, эколог, экономист, технолог, гидрогеолог — такая Междисциплинарность наряду со сложной структурой нефтегазовой отрасли определяют сложность процесса проектирования.

В процессе проектирования создается ряд моделей (геологическая, гидродинамическая, экономическая, технологическая), которые по сути своей представляют модели данных, а процесс моделирования можно рассматривать как последовательное повышение уровня и детализации модели (рис. 4) за счет вовлечения новых данных, а также применения различных алгоритмов интерполяции. Такое повышение уровня модели ведет к росту вычислительной нагрузки [9, 10].

Применение уровневого моделирования на основе комплексного подхода позволяет использовать модель достаточного уровня в зависимости от изученности, что решает ряд проблем:

- разрозненность моделей — разные типы моделей используются на разных этапах проектирования,



Рис. 1: Модель месторождения

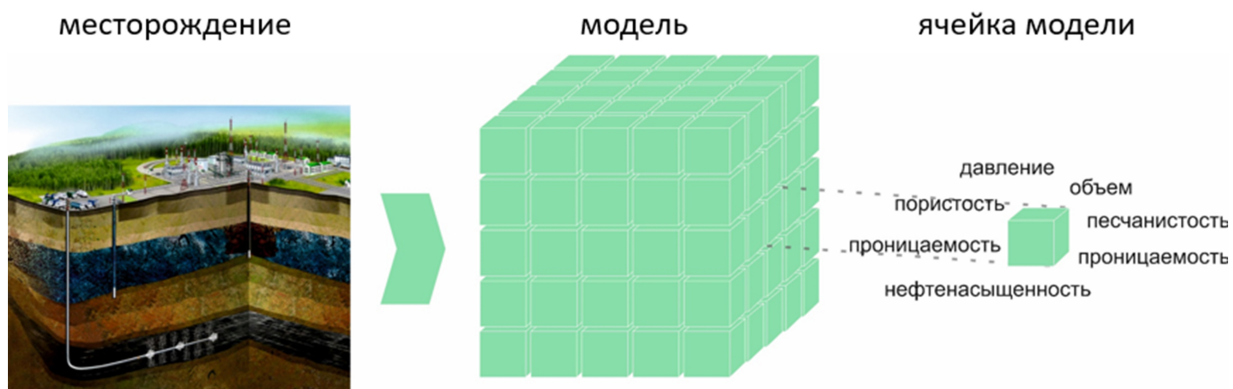


Рис. 3: Модель месторождения

- накопление ошибок — особенность технологии проектирования заключается в том, что моделирование выполняется последовательно, и создаваемые модели становятся основой для моделирования на следующих этапах,
- избыточность в моделировании — традиционно разработчик стремится к максимальной детализации модели, что в большинстве задач неоправданно,
- большое время расчета моделей и требовательность к вычислительным ресурсам. Предлагаемый комплексный подход был применен при проектировании разработки 14 месторождений Том-

ской области, находящихся на различных стадиях жизненного цикла, и позволил учесть риски разработки на стадии проектирования, а также снизить вычислительную нагрузку при выполнении задач оперативного управления и время выполнения проектов на 10-15%.

Заключение

Проектировщик обязан выполнять проектирование разработки руководствуясь регламентами [5, 6], которые устанавливают обязательным правилом выполнение трехмерного моделирования месторожде-

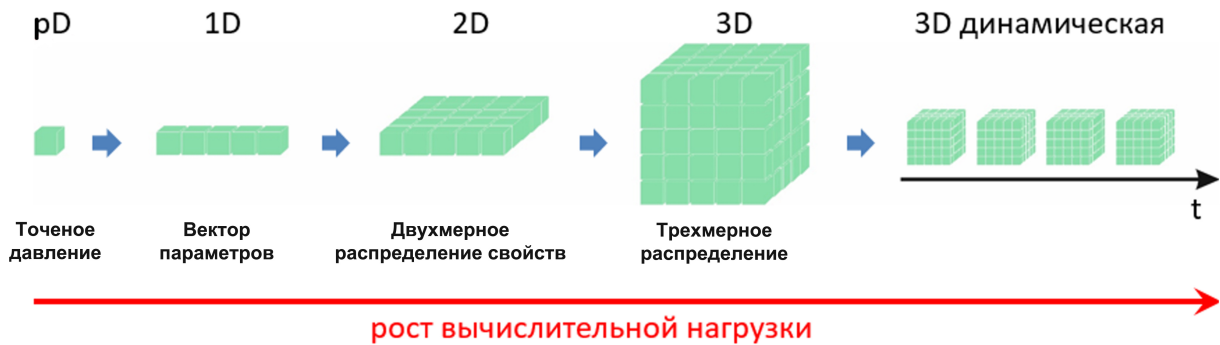


Рис. 4: Увеличение размерности модели

ний, однако они не устанавливают размерность или детализацию таких моделей.

Применение комплексного подхода позволяет дифференцированно выбирать мерность моделей в зависимости от задачи. Такой подход незаменим при выполнении проектирования разработки месторождений со слабой изученностью, например, для планирования геолого-технических мероприятий на отдельных участках с учетом их стоимости и рискованности их применения.

Работа выполнена при поддержке гранта Министерства образования и науки РФ № 2.1642.2017/4.6.

Литература

- [1] Zav'yalov D.A., Zakharova A.A., Shklyar A.V., Bagutdinov R.A. — An integrated approach to modeling by an example of a landfill of disposal of liquid oil waste // Программные системы и вычислительные методы. — 2017. — № 1. — С. 22-30.
- [2] Zakharova A., Vekhter E., Shklyar A., Zavyalov D. (2017) Visual Detection of Internal Patterns in the Empirical Data. In: Kravets A., Shcherbakov M., Kultsova M., Groumpos P. (eds) Creativity in Intelligent Technologies and Data Science. CIT&DS 2017. Communications in Computer and Information Science, vol 754.
- [3] Ямпольский В.З., Заикин И.А. Онтология «Интеллектуальное месторождение» // Известия Томского политехнического университета. — 2013 — Т. 323 — № 5 — С. 112-117.
- [4] Силич В.А., Комагоров В.П., Савельев А.О. Принципы разработки системы мониторинга и адаптивного управления разработкой «интеллектуального» месторождения на основе постоянно действующей геолого-технологической модели // Известия Томского политехнического университета. — 2013 — Т. 323 — №. 5. — С. 94-100.
- [5] Методические рекомендации по подготовке технических проектов разработки месторождений углеводородного сырья // Минприроды России. Утв. 18.05.2016.
- [6] Методические рекомендации по применению классификации запасов и ресурсов нефти и горючих газов // Минприроды России. Утв. 01.02.2016.
- [7] Бондарев А.Е., Галактионов В.А., Четкин В.М. Анализ развития концепций методов визуального представления данных в задачах вычислительной физики // Журнал вычислительной математики и математической физики. 2011. Т. 51. № 4. С. 669-683.
- [8] Mehrnoosh Oghbaie, Michael J. Pennock, William Rouse. Understanding the efficacy of interactive visualization for decision making for complex systems
- [9] Сазонов Е.О. Гидродинамическое моделирование процессов фильтрации плстовых флюидов при реализации потокоотклоняющих технологий // Нефтегазовое дело, 2013
- [10] Костюченко С.В., Бордзиловский А.С., Игнатов И.С., Шапиева Е.И. / Принцип итерационного сопряжения секторных моделей для полномасштабного моделирования больших и гигантских плстовых систем // Нефтяное хозяйство. — 2009. — №8. — С. 42-46.

Анализ производительности архитектуры ARM и другой перспективной архитектуры на SPEC тестах скомпилированных свободной коллекцией компиляторов GNU*

А. Ю. Агапов
 agapov_mipt@mail.ru

Московский физико-технический институт (государственный университет)

Данная работа посвящена анализу производительности двух микропроцессорных архитектур: ARM и другой перспективной архитектуры, имеющей схожую систему команд, названия которой пока нет, но условно будем обозначать ее SMI. В процессе разработки новой архитектуры требуется постоянно следить не только за корректностью исполнения на ней необходимых задач, но и по возможности отслеживать узкие места в системе команд для улучшения ее производительности. Измерение этой характеристики осуществляется на специальных программах-тестах, созданных корпорацией SPEC (Standart Performance Evaluation Corporation) [1]. SPEC имеет целью разработку и стандартизацию методов оценки производительности современных компьютеров. Разработанные SPEC тестовые пакеты являются де-факто стандартами для оценки производительности современных микропроцессоров, компьютеров и системного ПО.

В работе предлагается применение новой технологии для улучшения процесса отладки этих тестовых задач, а также осуществляется анализ эффективности системы команд ARM архитектуры с помощью технологии архитектурного профилирования.

Ключевые слова: архитектура ARM, архитектура на SPEC тестах, свободная коллекция компиляторов GNU.

Analysis of the performance of the ARM architecture and other perspective architecture in SPEC tests compiled by the free GNU compilers collection*

A. Yu. Agapov

Moscow Institute of Physics and Technology (State University)

This work is devoted to the analysis of the performance of two microprocessor architectures: ARM and another perspective architecture, which has a similar command system, the names of which are not yet available, but we will conditionally designate it as SMI. In the process of developing a new architecture, it is required to constantly monitor not only the correct execution of the necessary tasks on it, but also, if possible, monitor bottlenecks in the command system to improve its performance. Measurement of this characteristic is carried out on special test programs, created by the corporation SPEC (Standard Performance Evaluation Corporation) [1]. SPEC aims at developing and standardizing methods for evaluating the performance of modern computers. The SPEC test packs developed are de facto standards for evaluating the performance of modern microprocessors, computers and system software.

The paper proposes the use of new technology to improve the debugging of these test problems, and also analyzes the effectiveness of the ARM architecture command system using architectural profiling technology.

Key words: architecture ARM, architecture on SPEC tests, free collection of GNU compilers.

Введение

В процессе разработки перспективной архитектуры существует необходимость в исправлении некорректного исполнения задач на новой платформе. Решение этой проблемы, описано в первой главе этой работы. В ней предлагается применение достаточно эффективной технологии InterAspect для отладки программ, в которой реализована концепция аспектно-ориентированного про-

граммирования. С помощью терминов этой парадигмы в технологии описан удобный интерфейс языка Си, который скрывает от пользователя всю сложность машинного представления в процессе отладки задач, что существенно экономит время программиста.

ARM архитектура принадлежит семейству RISC, особенностью которого является тот факт, что все операции с памятью в этой системе команд выполняют определенные инструкции загрузки данных из памяти и сохранения в нее, называемые LDR и STR

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 16-07-00089 и 18-07-20030.

соответственно (от англ. «load register» и «store register»). Не трудно понять, что именно они являются самыми часто встречающимися в коде программы, поэтому требуют особого внимания в процессе оптимизации кода. Конечной целью этой работы было небольшое исследование этих инструкций и элементов их структур, с целью выявления их недостатков.

1. Методика отладки SPEC тестов

1.1. InterAspect инструментарий для GCC

GCC — это широко используемая, свободно распространяемая коллекция компиляторов, поддерживающая множество языков, таких как C, C++, Fortran, Java и Ada на более чем 30 целевых платформах. Это большая система программного обеспечения с более чем 100 разработчиками, вносящими свой вклад в проект на протяжении многих лет и с управляющим комитетом, состоящим из 13 экспертов, которые прилагают огромные усилия для сохранения архитектурной целостности. GCC преобразует каждый из его «front-end» языков на промежуточное, независимое от языка представление, называемое GIMPLE, которое затем транслируется в машинный код одной из многих целевых архитектур, поддерживаемых GCC.

Поддержка в GCC написания плагинов предоставляет хорошую возможность для разработки практичного, широко используемого инструментария для изменения программ с целью проверки «на лету», а так как плагины работают на уровне представления GIMPLE, то инструментарий применим для любого из «front-end» языков GCC. Системы преобразования, управляющие машинным кодом, так же могут работать с любым языком, но низкоуровневый машинный код намного тяжелее анализировать, и он предоставляет меньше информации, в отличие от GIMPLE.

Применение GCC плагин-инструментария предоставляет хорошее преимущество пе-

ред машинным кодом, но в тоже время имеет значительный недостаток: несмотря на то, что GIMPLE — это промежуточное представление, оно также является низкоуровневым языком, требующее построения абстрактного синтаксического дерева (Abstract Syntax Tree — AST) обходных функций, для того чтобы преобразовать одно GIMPLE выражение в другое. Таким образом, написание плагинов является непросто задачей даже для тех, кто близко знаком с GIMPLE спецификой.

Чтобы решить эту проблему, командой разработчиков [2] был создан, так называемый, InterAspect инструментарий, который позволяет разрабатывать плагины с помощью терминов аспектно-ориентированного программирования (АОП). Будучи примененным, InterAspect использует плагин-интерфейс GCC для управления GIMPLE, но скрывает всю его сложность от пользователя, заменяя его аспектно-ориентированным интерфейсом, в котором инструментарий оформлен в качестве, так называемых «срезов» (от англ. «roincuts»). Срезы — это набор точек вхождения в программу, именуемые «точками соединения» (от англ. «join points»), в которых выполняются «функции совета» (или просто «советы», от англ. «advices»), встраиваемые в процессе, так называемой «прошивки» (от англ. «weaving»).

Интерфейс InterAspect позволяет пользователю модифицировать процесс прошивки, определяя функции обратного вызова (от англ. «callback functions») в каждой точке соединения, и имеющие доступ к определенной информации в них. Эти функции могут использовать эту информацию и результаты статического анализа для модификации встраиваемого инструментария.

Таким образом, InterAspect предоставляет следующие функциональные возможности:

1. InterAspect построен на основе GCC — компилятора, имеющего армию преданных поклонников.

2. InterAspect предоставляет интерфейс, который упрощает и улучшает процесс проверки открытого кода.
3. InterAspect имеет доступ к внутренней структуре GCC, что позволяет использовать статический анализ и методы метапрограммирования в процессе прошивки.

1.2. Обзор структуры GCC и InterAspect

Как показано на рис. 1, GCC транслирует каждый из его «front-end» языков в промежуточное представление GIMPLE для анализа и оптимизации.

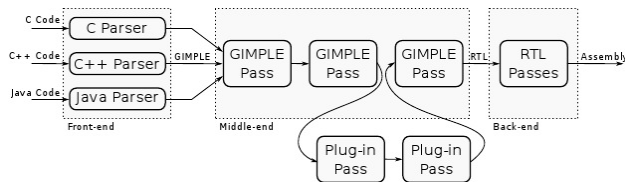


Рис. 1: Упрощенная схема GCC компиляции

Каждое преобразование GIMPLE кода определяет его собственную фазу. Эти фазы (некоторые из них могут быть плагинами) образуют промежуточный уровень компиляции GCC — middle-end. Более того, плагин-фаза может быть реализована с помощью InterAspect, позволяя плагину добавлять инструментарий прямо в GIMPLE код. Затем, самая нижняя оптимизированная и инструментированная фаза GIMPLE преобразуется в уровень регистровых передач (Register Transfer Level, RTL), который в конечном итоге транслируется в ассемблер.

Код GIMPLE похож на код Си, но при этом все операции имеют не более трёх операндов. Сложные выражения (возможно с побочными эффектами) путём введения временных переменных разбиваются на простые трёхадресные, а информация о типах сохраняется для каждого операнда в каждом GIMPLE выражении.

Рис. 2 показывает Си программу и соответствующий ей GIMPLE код, который сохраняет всю информацию на уровне исходного

кода, такую как тип данных и процедурные вызовы. Хотя в примере не показано, но типы данных GIMPLE также включают в себя указатели и структуры.

```

int main() {
  int a, b, c;
  a = 5;
  b = a + 10;
  c = b + foo(a, b);
  if (a > b + c)
    c = b++ / a + (b * a);
  bar(a, b, c); }
=>
1. int main {
2.   int a, b, c;
3.   int T1, T2, T3, T4;
4.   a = 5;
5.   b = a + 10;
6.   T1 = foo(a, b);
7.   c = b + T1;
8.   T2 = b + c;
9.   if (a <= T2) goto fi;
10.  T3 = b / a;
11.  T4 = b * a;
12.  c = T3 + T4;
13.  b = b + 1;
14.  fi: bar (a, b, c); }
  
```

Рис. 2: Простая Си программа и соответствующий ей GIMPLE код

Недостатком работы с GIMPLE представлением является невидимость некоторых специфических языковых конструкций в его коде. К примеру, указание в качестве среза определенных типов циклов в настоящее время является невозможным, т. к. все циклы в GIMPLE коде выглядят одинаково. InterAspect может быть расширен языковыми срезами, имплементированными на базе AST.

Структура InterAspect

InterAspect инструментарий работает следующим образом: он встраивает в компиционный процесс фазу прохождения через GIMPLE код для определения программных точек, которые являются точками соединения в конкретном срезе. Для каждой такой точки соединения он вызывает предоставленные пользователю функции обратного вызова, которые встраивают обращение к функциям совета. Эти функции могут быть написаны на любом языке, который будет связан с целевой программой, и могут модифицировать ее и глобальные переменные. Совет нужный для поддержки дополнительного состояния может объявлять как статические, так и глобальные переменные. В отличие от традиционных АО систем, которые используют специальный АО язык для определения срезов, InterAspect предоставляет интерфейс языка Си для этих целей.

Как показано на рис. 3, InterAspect также может служить как «back-end» инструментарий для традиционного АО языка спецификаций. Задача такого языкового компилятора состоит в разбиении функциональности АОП на три задачи: определение срезов, определение связанных с ними инструкций прошивки и функций совета. Первые две задачи выполняются модулем прошивки для вычислений в процессе плагин фазы в то время, как код функций совета подается на вход GCC для компиляции.

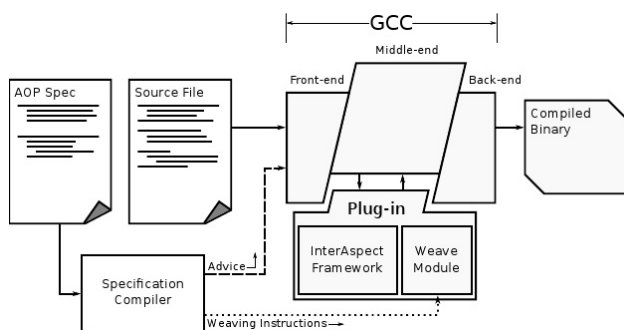


Рис. 3: Структура InterAspect инструментария для GCC

1.3. InterAspect интерфейс

Эта глава описывает функции интерфейса InterAspect инструментария, большинство из которых делятся на две категории: (1) функции для создания и фильтрации срезов и (2) функции для определения и модификации точек соединения. Учтите, что, используя их, пользователи InterAspect могут писать плагины сами, причем для этого не нужно включать какие-либо заголовочные файлы GCC или управлять непосредственно структурой данных.

Создание и фильтрация срезов

Первым шагом для добавления инструментария является создание срезов, которые используют, так называемые функции совпадения. Текущая версия InterAspect инструментария поддерживает 4 функции, представленных в таблице 1, позволяя создавать 4 типа срезов.

Первые два типа — это срез на вхождение и выход из функции — делают воз-

можным добавление инструментария, который запускается при каждом ее выполнении. Третий тип — срез на вызов функции — может определить точки обращения к ней, а также перехватывать вызовы к библиотечным функциям без их перекомпиляции. Последний тип — срез на присваивание — удобен для отслеживания изменений значений переменных.

Таблица 1: Функции совпадения для создания срезов

<code>struct aop_pointcut *aop_match_function_entry()</code>	Создает срез, определяющий каждую точку вхождения в функцию
<code>struct aop_pointcut *aop_match_function_exit()</code>	Создает срез, определяющий каждую точку выхода из функции
<code>struct aop_pointcut *aop_match_function_call()</code>	Создает срез, определяющий каждый вызов функции
<code>struct aop_pointcut *aop_match_assignment_by_type(struct aop_type *type)</code>	Создает срез, определяющий каждое присваивание значения переменной или запись в память, соответствующее определенному типу

После определения функций совпадения, плагин может детализировать их с помощью функций фильтрации. Эти функции добавляют дополнительные ограничения на срез, убирая точки соединения, которые им не удовлетворяют. К примеру, возможна фильтрация среза по вызову функций, которые возвращают определенный тип переменных или по вызову к конкретной функции (см. таблицу 2).

Таблица 2: Функции фильтрации для детализации срезов на вызов функции

<code>void aop_filter_call_pc_by_name(struct aop_pointcut *pc, const char *name)</code>	Фильтрует вызовы функций с данным именем
<code>void aop_filter_call_pc_by_param_type(struct aop_pointcut *pc, int n, struct aop_type *type)</code>	Фильтрует вызовы функций, у которых n-ый аргумент совпадает с заданным типом
<code>void aop_filter_call_pc_by_return_type(struct aop_pointcut *pc, struct aop_type *type)</code>	Фильтрует вызовы функций по возвращаемому ими типу данных

Реализация точек соединения

InterAspect плагины обходят точки соединения среза, вызывая для каждого сре-

за определённую разработчиком функцию `join` (см. таблицу 3). Затем `InterAspect` вызывает эту функцию для каждой точки соединения, инструментируя их вызовом функции совета.

Функции обратного вызова используют так называемые функции захвата для проверки значений, связанных с точкой соединения. Они воздействуют на два типа переменных: статические, значения которых становятся известным в процессе компиляции, и динамические, чье содержимое проявляется только при исполнении программы.

Таблица 3: Определение функции соединения для выполнения итераций над срезом

<pre>void aop_join_on(struct aop_pointcut *pc, join_callback callback, void *callback_param)</pre> <p>Поддерживает обратную связь с любой структурой данных объявленных в поле <code>callback_param</code></p>
--

Неизменяемые значения, такие как имя присваиваемой переменной в операторе присваивания, могут быть прочитаны непосредственно из обратного вызова, в отличие от динамических — их можно передать в качестве параметров функциям совета. Таким образом, доступ к ним осуществляется во время запуска программы. Эти динамические переменные представлены в функциях обратного вызова специальными `aop_dynval` объектами. Функции захвата специфицируются по типам точек соединения, над которыми они оперируют (см. таблицы 4 и 5).

Таблица 4: Функции захвата для точек соединения типа «вызов-функции»

<pre>const char *aop_capture_lhs_name(aop_joinpoint *jp)</pre> <p>Перехватывает имя функции, вызванной в данной точке соединения</p>
<pre>struct aop_dynval *aop_capture_param(aop_joinpoint *jp, int n)</pre> <p>Перехватывает значение n-ого параметра, переданного функции в данной точке соединения</p>
<pre>struct aop_dynval *aop_capture_return_value(aop_joinpoint *jp)</pre> <p>Перехватывает значение, возвращаемое функцией в данной точке соединения</p>

Таблица 5: Функции захвата для точек соединения типа «присваивание»

<pre>const char *aop_capture_lhs_name(aop_joinpoint *jp)</pre> <p>Перехватывает имя переменной, присваиваемой в данной точке соединения типа «присваивание» или возвращает <code>NULL</code>, если точка соединения не является присвоением именованной переменной</p>
<pre>enum aop_scope aop_capture_lhs_var_scope(aop_joinpoint *jp)</pre> <p>Перехватывает диапазон переменной, присвоенной в данной точке соединения типа «присваивание». Переменные могут иметь глобальный, локальный в пределах файла и локальный в пределах функции диапазон. Если точка соединения не является присвоением переменной, эта функция возвращает <code>AOP_MEMORY_SCOPE</code></p>
<pre>struct aop_dynval *aop_capture_lhs_addr(aop_joinpoint *jp)</pre> <p>Перехватывает адрес памяти, присвоенный в данной точке соединения типа «присваивание»</p>
<pre>struct aop_dynval *aop_capture_assigned_value(aop_joinpoint *jp)</pre> <p>Перехватывает присвоенное значение в данной точке соединения типа «присваивание»</p>

Аспектно-ориентированные системы, такие как `AspectJ` [3] предоставляют использование логических операторов «И» и «ИЛИ». `InterAspect` интерфейс также может быть расширен ими, но без них можно обойтись. Того же самого результата можно добиться в `InterAspect` инструментарии, включая соответствующую логику в обратный вызов. К примеру, плагин способен инструментировать вызовы к функциям `malloc` и `free`, объединяя их в один срез со всеми вызовами функций и используя средство `aop_capture_function_name`, чтобы добавить обращение функции совета только к `malloc` и `free`. Более того, такими простыми случаями, как этот, можно управлять с помощью регулярных выражений для проверки совпадений имени функции с заданным. Эта функциональность будет скоро добавлена в оболочку.

После перехвата, обратный вызов может добавить обращение к функции совета до или после точки соединения, используя функции вставки (см. таблицу 6). На вход функции `aop_insert_advice` во время запуска программы подается любое количество параметров для передачи их функции совета, включая значения, перехваченные из точки соединения и значения, вычисленные плагином, как таковым, в процессе оснащения инструментальными средствами.

Таблица 6: Функция вставки для модифицирования точки соединения с помощью вызова функции совета

```
void aop_insert_advice(struct aop_joinpoint *jp, const char
*advice_func_name, enum aop_insert_location location, ...)
```

Встраивает вызов совета до или после точки соединения (зависит от значения переменной location), передавая любое количество параметров. Плагин получает точку соединения, проходя по срезу, с помощью функции aop_join_on

Использование обратного вызова для прохождения по конкретной точке соединения делает возможным модификацию инструментария в каждом рассматриваемом узле. Плагин может перехватить значения около точки соединения, с целью решить, какую из функций совета вызвать, какой параметр ей передать или добавлять ли вообще этот совет.

Дублирование функций

InterAspect предоставляет возможность дублирования функций для переключения инструментария на функциональном уровне. Несмотря на то, что добавление советов на GIMPLE уровне является очень эффективным средством анализа кода, пользователь по разным причинам может захотеть переключаться между инструментированным и неинструментированным кодом. Дублирование создает две или более копии тела функции (которые позже могут быть модифицированы инструментарием по-разному) и переопределяет ее, вызывая перед входом в нее специальный совет, который определяет какую из копий функции запустить.

При включении в срез точки соединения, являющейся вызовом функции с дублированным телом, пользователь определяет, какая из копий должна быть добавлена в срез. Плагин может создать функцию, инструментарий которой может быть включен или выключен во время запуска, просто добавив набор инструментов в одну из копий ее тела. Также существует возможность переключаться между разными видами инструментария.

1.4. Применение технологии InterAspect к отладке SPEC тестов

1.4.1. Визуализация структуры данных «куча»

Визуализатор кучи использует InterAspect интерфейс для наглядного просмотра событий, происходящих в памяти. Он может сгенерировать графическое представление кучи в реальном времени в течение исполнения программы. Аллоцированные объекты представляются прямоугольниками, переменные-указатели и поля — овалами, а ребра графа указывают, куда они ссылаются.

Для того чтобы нарисовать граф, визуализатору необходимо перехватить объекты распределения и освобождения памяти и присваивания указателей, которые меняют ребра графа. Рис. 4 показывает прототип визуализатора, использующего Graphviz [4] — свободно распространяемую программу для рисования графов. На изображении видно три узла в связанном списке в процессе сортировки методом пузырька. Каждый узел помечен его размером, адресом памяти и адресами его полей. Переменные, указывающие на NULL, или на неверное расположение памяти помечены пунктирной границей. Ребра графа помечены номером строки в файле, которой было создано это ребро, а также количеством присваиваний к исходной переменной.

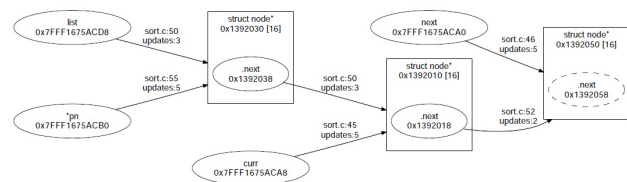


Рис. 4: Визуализация «кучи» в процессе сортировки методом пузырька в связанном списке

Код InterAspect для визуализатора кучи инструментирует каждое размещение в памяти (вызов функции malloc) вызовом функции совета heap_allocation, а каждое присваивание указателя — вызовом функции pointer_assign. Эти функции совета об-

новляют граф. Управление инструментированием других функций размещения и освобождения памяти, к примеру, таких как `calloc` и `free`, осуществляется аналогично.

Код `InterAspect`, показанный на рис. 5, с помощью совета `instrument_malloc_calls` создает срез на вызов функции `malloc`, а затем обращается к `aop_join_on` для прохождения по всем точкам ее вызова в срезе. И лишь короткая функция `main` (на рисунке не показана) нужна для того, чтобы `GCC` запустил функцию `instrument_malloc_calls` в процессе компиляции. Функция `aop_match_function_call` конструирует исходный срез, включающий в себя обращения ко всем функциям, а дополнительные фильтры сводят срез к включению только самой `malloc`. Первым стоит фильтр `aop_filter_pc_by_name`, который отбрасывает обращения к функциям, чье имя не совпадает с `malloc`. Затем фильтры `aop_filter_pc_by_param_type` и `aop_filter_pc_by_return_type`, отсеивают функции, несовпадающие со стандартным прототипом `malloc`, который на вход в качестве параметра принимает целое беззнаковое число и возвращает указатель на адрес. Этот отсеивающий шаг необходим, т. к. программа может определять свою собственную функцию `malloc` с другим прототипом. Для каждой точки соединения в срезе (в данном случае это выражение, вызывающее функцию `malloc`) функция `aop_join_on` обращается к `malloc_callback`. Два запуска функций захвата в теле обратного вызова возвращают объекты типа `aop_dynval` для первого вызываемого параметра, а также следующие значения: размер выделенной памяти и ее адрес соответственно. `Aop_dynval` объекты служат в качестве структурного нуля для тех значений, которые не могут быть вычислены в процессе компиляции.

В конечном итоге `aop_insert_advice`, принимая в качестве входных параметров два захваченных значения, добавляет обращение к функции совета. Обратите внима-

```
static void instrument_malloc_calls()
{
    /* Конструирование среза, отслеживающего обращения к функции: void *malloc(unsigned int). */
    struct aop_pointcut *pc = aop_match_function_call();
    aop_filter_call_pc_by_name(pc, "malloc");
    aop_filter_call_pc_by_param_type(pc, 0, aop_t_all_unsigned());
    aop_filter_call_pc_by_return_type(pc, aop_t_all_pointer());

    /* Прохождение по всем выражениям в срезе. */
    aop_join_on(pc, malloc_callback, NULL);
}

/* Функция malloc_callback() единожды исполняется для каждого вызова функции
malloc(), инструментируя каждое обращение к ней вызовом функции heap_allocation(). */
static void malloc_callback(struct aop_joinpoint *jp, void *arg)
{
    struct aop_dynval *object_size;
    struct aop_dynval *object_addr;

    /* Захват размера выделяемого объекта и адреса его размещения в памяти. */
    object_size = aop_capture_param(jp, 0);
    object_addr = aop_capture_return_value(jp);

    /* Добавление обращения к функции совета, считывая размер и адрес объекта в качестве входных
    параметров. (Параметр AOP_TERM_ARG необходим для указания конца списка аргументов. Это
    связано с погрешкой работы varargs функций языка C.) */
    aop_insert_advice(jp, "heap_allocation", AOP_INSERT_AFTER,
        AOP_DYNVAL(object_size), AOP_DYNVAL(object_addr),
        AOP_TERM_ARG);
}
```

Рис. 5: Инструментирование всех событий, связанных с распределением памяти

ние, что `InterAspect` выбирает тип для своих значений, основываясь на том, как они были отфильтрованы. В данном случае использованные здесь фильтры сужают множество типов, которым могут принадлежать переменные `object_size` и `object_addr` до целого беззнакового и некоторого типа указателя соответственно. Таким образом, `InterAspect` предполагает, что функция совета `heap_allocation` имеет следующий прототип:

```
void heap_allocation (unsigned long long,
                    void*)
```

Для поддержки этой функциональности `InterAspect` код во время запуска программы должен отфильтровывать переменные по типу для их дальнейшего захвата и использования.

`InterAspect` код на рис. 6 отслеживает присваивания указателю следующего вида:

```
list_node → next = new_node,
```

где функция `aop_match_assignment_by_type` создает срез, удовлетворяющий определенному типу присваиваний. В данном случае

мы интересуемся только присваиваниями к указателю.

```
static void instrument_pointer_assignments()
{
    /* Конструирует срез отлавливающий все присваивания к указателю. */
    struct aop_pointcut *pc = aop_match_assignment_by_type(aop_t_all_pointer());

    /* Прохождение по всем выражениям в срезе. */
    aop_join_on(pc, assignment_callback, NULL);
}

/* Функция assignment_callback единожды выполняется для каждого присваивания к указателю и
инструментирует их обращением к функции pointer_assign(). */
static void assignment_callback(struct aop_joinpoint *jp, void *arg)
{
    struct aop_dynval *address;
    struct aop_dynval *pointer;

    /* Перехватывает значение указателя и адрес, куда оно кладется. */
    address = aop_capture_lhs_addr(jp);
    pointer = aop_capture_assigned_value(jp);
    aop_insert_advice(jp, "pointer_assign", AOP_INSERT_AFTER,
        AOP_DYNVAL(address), AOP_DYNVAL(pointer),
        AOP_TERM_ARG);
}
```

Рис. 6: Инструментирование всех присваиваний к указателям

Для каждой точки соединения, содержащей присваивание, функция `assignment_callback` захватывает значение указателя и адрес, по которому оно кладется, и сохраняет их в переменные `pointer` и `address`, эти значения используются визуализатором кучи для определения конца и начала ребра графа соответственно. В данном случае это будут значения `new_node` и `&list_node` → `text`. Функция `aop_capture_lhs_addr`, захватывающая `address` не требует подобной фильтрации по типу перехватываемых переменных, т.к. адрес всегда принадлежит типу указателей.

Значение, отловленное функцией `aop_capture_assigned_value`, и сохраненное в переменной `pointer` имеет тип `void`, т.к. мы фильтровали срез только по присваиваниям к указателям. В результате `InterAspect` подразумевает, что функция совета `pointer_assign` имеет следующий прототип:

```
void pointer_assign(void*, void*).
```

1.4.2. Анализатор диапазона целочисленных переменных

Это приложение создано с целью выявления аномального поведения программы

посредством отслеживания границ значений целочисленных переменных. Программа может определить нормальный диапазон значений из пробных запусков программы с «хорошими» входными данными. Значения, выходящие за его пределы, в будущих запусках помечаются как аномальные, идентифицируя место возможной неполадки. К примеру, неправильное значение переменной, используемое в качестве номера элемента массива, может указывать за его пределы, вызывая тем самым ошибку.

Анализатор использует выборочный метод для уменьшения накладных расходов на инструментирование программы, из-за этого теряется точность определения диапазона, но выбор этого метода оправдан в большинстве случаев. Выборка может быть реализована как случайно, так и способом, предложенным в [5].

`InterAspect` умеет дублировать тело функции с целью поддержки функциональности включения и выключения добавленного в них инструментария. В этом случае создается две копии тела функции, в одну из которых плагин может добавить совет, а распределяющий блок в ее начале решает какую из копий запустить.

Рис. 7 показывает, как можно использовать `InterAspect` для отслеживания изменений в целочисленных переменных. Обращение к `aop_duplicate` делает копию тела каждой функции. Первый ее аргумент говорит о том, что копий функции должно быть две, а второй указывает на имя функции, которую должен вызвать распределительный блок, с целью решить, какую из копий запускать. Во время запуска дубликата функции распределитель вызывает `distributor_func`, которая должна возвращать целое число. Копии нумеруются, начиная с нуля, и `distributor_func` возвращает значение, определяющее какой из этих копий распределительный блок должен передать управление.

Вместо функции `aop_join_on` в данном случае используется `aop_join_on_coru`, ко-

```

static void instrument_integer_assignments()
{
    struct aop_pointcut *pc;

    /* Дублирование тела функции. */
    aop_duplicate(2, "distributor_func", AOP_TERM_ARG);

    /* Построение среза, отлавливающего все присваивания целого типа. */
    pc = aop_match_assignment_by_type(aop_t_all_signed_integer());

    /* Прохождение по всем выражениям в срезе */
    aop_join_on_copy(pc, 1, assignment_callback, NULL);
}

/* Функция assignment_callback инструментрует каждое присваивание целого типа обращением к
функции int_assign(). */
static void assignment_callback(struct aop_joinpoint *jp, void *arg)
{
    const char *variable_name;
    int variable_index;
    struct aop_dynval *value;
    enum aop_scope scope;
    variable_name = aop_capture_lhs_name(jp);
    if (variable_name != NULL) {
        /* Выбор номера индекса для данной переменной. */
        scope = aop_capture_lhs_var_scope(jp);
        variable_index = get_index_from_name(variable_name, scope);
        aop_insert_advice(jp, "int_assign", AOP_INSERT_AFTER,
            AOP_INT_CST(variable_index), AOP_DYNVAL(value),
            AOP_TERM_ARG);
    }
}

```

Рис. 7: Отслеживание изменений в целочисленных переменных

торая проходит по всем точкам соединения данной копии функции. В конечном итоге получается, что только один из дубликатов функции инструментирован, а второй остается нетронутым.

Функция обратного вызова идентична тем, которые мы использовали в разделе 1.4.1. Главное отличие — дополнительное обращение к `get_index_from_name`, которая преобразует имя переменной в целочисленный индекс. Эта функция (здесь не показана для краткости) также определяет диапазон переменной таким образом, что она может присвоить разные индексы локальным переменным в разных функциях. В дальнейшем будет возможно напрямую подавать на вход имя переменной в качестве входного параметра функции совета, но постоянный поиск переменной по ее имени увеличит временные затраты.

Функция `aop_capture_lhs_name` возвращает строку вместо `aop_dynval` объекта, потому что имена переменных известны во время компиляции. Необходимо проверять возвращаемое значение на `NULL`, т. к. не

все присваивания осуществляются именованным переменным.

Для того чтобы лучше понять и оценить временные издержки, этот плагин был протестирован на утилите для сжатия данных `bzip2`, используя пустую функцию совета. Инструментированная программа `bzip2` содержала обращения к советам в каждом операторе целочисленного присваивания, но сами советы являлись пустыми функциями, что позволило измерить накладные расходы только на их вызовы. Вместе с распределительным блоком, увеличивающим временные издержки, из-за постоянного выбора инструментированного тела функции было получено 24% замедление. Дублирование функций само по себе вносит очень маленький вклад в накладные расходы, то есть в том случае, если распределительный блок выбирает всегда не инструментированное тело функции, программа замедляется незначительно.

1.4.3. Покрытие кода

Самый простой путь измерения покрытия кода — это выбрать определенным образом срез и измерить процентное соотношение точек соединения, исполняемых в процессе работы программы. С помощью функции прохождения по всем точкам соединения, можно легко их пометить и отслеживать в процессе исполнения программы.

Пример на рис. 8 добавляет инструментарий для отслеживания покрытия точек вхождения и выхода из функции.

Для уменьшения накладных расходов функция `choose_unique_index` присваивает целочисленный индекс каждой отслеживаемой точке соединения схожим образом, как и в главе 1.4.2. Каждый индекс сохраняется вместе с соответствующим именем файла и номером строки в нем с помощью функции `save_index_to_disk`. Во время запуска программы выполняется совет, который выдает только последовательность покрытых индексов, а специальная утилита после выполнения программы использует эти

```

static void instrument_function_entry_exit()
{
    struct aop_pointcut *entry_pc;
    struct aop_pointcut *exit_pc;

    /* Конструирование двух срезов: один на вход в функцию, другой на выход из нее. */
    entry_pc = aop_match_function_entry();
    exit_pc = aop_match_function_exit();
    aop_join_on(entry_pc, entry_exit_callback, NULL);
    aop_join_on(exit_pc, entry_exit_callback, NULL);
}

/* Функция entry_exit_callback присваивает индекс каждой точке соединения, которую видит и
сохраняет этот индекс на диск. */

static void entry_exit_callback(struct aop_joinpoint *jp, void *arg)
{
    int index, line_number;
    const char *filename;
    index = choose_unique_index();
    filename = aop_capture_filename(jp);
    line_number = aop_capture_lineno(jp);
    save_index_to_disk(index, filename, line_number);
    aop_insert_advice(jp, "mark_as_covered", AOP_INSERT_BEFORE,
        AOP_INT_CST(index), AOP_TERM_ARG);
}

```

Рис. 8: Инструментирование входа в функцию и выхода из нее

данные для вычисления процентного соотношения покрытых точек соединения.

2. Методика оценки производительности архитектуры при помощи SPEC тестов

2.1. Основные определения

SPEC (от англ. Standard Performance Evaluation Corporation) — корпорация по стандартизации оценки производительности, является некоммерческой организацией, членами которой являются производители компьютерной техники, компании по производству программного обеспечения, университеты, исследовательские организации, издатели и консультанты. Цель SPEC в том, чтобы создавать, поддерживать и одобрять стандартный набор соответствующих критериев для компьютерных систем. Хотя ни один набор тестов не может в полной мере охарактеризовать общую производительность системы, SPEC считает, что сообществу пользователей может быть полезен набор объективных тестов, которые могут служить в качестве общей точки отсчета.

Бенчмарк, тест производительности (англ. benchmark) — контрольная задача, необходимая для определения сравнительных ха-

рактеристик производительности компьютерной системы. Как правило, бенчмарк — это компьютерная программа, которая выполняет строго определенный набор операций, называемый нагрузкой и возвращает результат в определенной форме, являющейся метрикой, т. е. оценкой того, как хорошо выполнялась программа на данном компьютере. Обычно в качестве метрики берут время выполнения программы или пропускную способность — сколько единиц нагрузки были завершены в единицу времени.

Успешное прохождение ряда бенчмарков является свидетельством стабильности системы в штатном и в форсированном режимах, поэтому бенчмарки используются для сравнения производительности компьютеров и часто являются критерием для выбора компонента того или иного производителя.

SPEC CPU2006 — пакет тестов, фокусирующийся на ресурсоемких вычислениях, это означает, что эти тесты зависят от производительности:

1. Процессора компьютера.
2. Архитектуры памяти.
3. Компилятора.

Необходимо помнить о последних двух компонентах. Не следует думать, что производительность SPEC CPU тестов зависит только от характеристик процессора. SPEC CPU2006 состоит из двух наборов программ, которые фокусируются на двух разных типах ресурсоемких вычислений:

- CINT2006 — измеряет производительность целочисленных операций;
- CFP2006 — измеряет производительность операций с плавающей запятой.

Каждый набор может быть использован для измерения производительности по следующим направлениям:

- Компиляционный метод: Использует для сборки всех тестов одни и те же совместимые с компилятором опции (base метрика), либо для каждого бенчмарка

подбирает опции индивидуально (peak метрика).

- Скорость или пропускная способность: Использует в качестве метрики время выполнения задач (speed метрика) или суммарную производительность набора одновременно запущенных задач (rate метрика).

SPEC CPU2006 не предназначен для нагрузки других компонент вычислительной машины, таких как: работа в сети, операционная система, графика или система ввода/вывода. Производительность бенчмарков, предназначенных для тестирования одного ядра, практически не зависит от этих компонентов, но производительность больших rate-задач может быть сильно искажена влиянием операционной системы или системой ввода/вывода.

2.2. Метрики производительности

Как уже отмечалось выше, тестовые задачи SPEC CPU2006 разделены на две большие подгруппы — целочисленных задач CINT2006 и задач с плавающей точкой CFP2006. Соответственно, тестовый пакет SPEC CPU2006 предусматривает возможность получения двух фундаментальных оценок производительности, в общем виде соответственно именуемых SPECint2006 и SPECfp2006.

Для каждой из этих оценок предусмотрены две важнейшие метрики производительности системы. Первая из них измеряет, насколько быстро (т. е. за какой промежуток времени) система способна решить одну задачу — эта метрика называется метрикой скорости (speed). Вторая из них отражает, какое количество задач способна решить система в определенный, фиксированный промежуток времени. Эта метрика именуется метрикой пропускной способности (throughput, или rate). В первом случае осуществляется запуск одной копии задачи (при этом сама задача может быть автоматически распараллелена оптимизирующим компилятором с целью использо-

вания всех доступных системных процессоров — такое измерение по-прежнему будет являться метрикой скорости). Во втором случае производится одновременный параллельный запуск нескольких копий одной и той же задачи, количество которых обычно устанавливается равным количеству системных процессоров, но это не является абсолютным требованием метрики пропускной способности.

Наконец, в зависимости от типовых сценариев использования компиляторов программистами, есть два вида пользователей:

1. Некоторые пользователи для достижения наилучшей производительности могут экспериментировать с разными компиляторами и опциями, а также разрабатывать свой собственный многопроходный процесс сборки тестов или свой собственный набор входных данных.
2. Другие пользователи предпочитают относительную простоту, использование небольшого количества опций и однопроходный процесс сборки.

Вдобавок вышесказанному, существует огромное количество других моделей использования тестов, варьирующиеся от простой компиляции без всяких опций оптимизации до встраивания собственных директив и изменения исходного кода. Возникает резонный вопрос, какие из вышеперечисленных моделей использования бенчмарков CPU2006, корпорация SPEC может разрешить, а какие нет?

Также ясно, что для пользователей, преследующих иные цели, выбор из этого бесконечного числа моделей может быть сделан абсолютно произвольно. Тем не менее, выбор должен быть сделан. SPEC решила каждую из метрик speed и rate измерять в двух различных вариантах компиляции кода, которые получили название базового (base) и пикового (peak) вариантов:

1. Базовая метрика или base, например SPECint_base2006, является обязательной при отправке результатов измерения производительности системы на

сайт корпорации SPEC и имеет строгие ограничения к компиляционному процессу. К примеру, все бенчмарки, написанные на одном программном языке, должны использовать одни и те же опции компиляции в строго заданном порядке. Эта метрика подходит для тех, кто предпочитает относительно простой процесс сборки.

2. Пиковая метрика или peak, например SPECint 2006, является опциональной и имеет меньшее количество строгих ограничений и отражает более «экспериментальный» подход программиста к компиляции кода. Она позволяет использовать разные опции компиляции для каждого бенчмарка, а также оптимизации, управляемые обратной связью. Это подходит для желающих потратить больше времени и усилий на процесс сборки тестов.

Следует знать относительно последней метрики, что многопроходная компиляция кода относится именно к пиковой метрике и недопустима при сборке задач с целью получения «базового» результата (что допускалось в предыдущей версии SPEC CPU2000).

3. Анализ возможности улучшения производительности архитектуры

Данное исследование было проделано на начальной стадии разработки новой перспективной архитектуры SMI. Ее система команд имеет схожую структуру с архитектурой ARM, поэтому возникает большой интерес в исследовании узких мест системы команд ARM для их учета в процессе дизайна платформы SMI. Конечной целью данной работы является исследование эффективности последней версии системы команд ARMv7, поэтому рассмотрим кратко основные ее особенности.

3.1. Краткое описание системы команд ARMv7

3.1.1. Основная идея архитектуры ARM

Целью компании ARM является разработка RISC архитектур с набором команд для эффективного исполнения специфических приложений с оптимальным балансом между программной и аппаратной реализацией. ARM архитектура поддерживает реализации в широком диапазоне требований по производительности. Она проявила себя в качестве доминантной во многих сегментах рынка. Простота реализации ARM процессоров приводит к конечным продуктам малых размеров, а малые размеры, в свою очередь, к очень низкому энергопотреблению. Малые размеры, производительность и низкое энергопотребление являются ключевыми атрибутами ARM платформы.

ARM — это архитектура семейства RISC (Reduced Instruction Set Computer), поэтому включает в себя следующие свойства, типичные для этого семейства:

- единый большой регистровый файл;
- load/store архитектура, в которой операции с памятью осуществляются только над содержимым регистров, а не напрямую с ячейками памяти;
- простые режимы адресации, со всеми load/store адресами, определенными только из содержимого регистров и полей инструкции.

Вдобавок ко всему вышесказанному ARM архитектура предоставляет:

- инструкции, которые комбинируют побитовый сдвиг с арифметическими или логическими операциями;
- авто-инкрементный и авто-декрементный режимы адресации для оптимизации программных циклов;
- инструкции, поддерживающие загрузку из памяти и сохранение в память содержимого более одного регистра, для увеличения пропускной способности работы с памятью;
- условное исполнение почти у всех инструкций для увеличения пропускной способности исполнения кода.

Эти улучшения базовой RISC архитектуры позволяют ARM процессорам достичь оп-

тимального баланса между высокой производительностью, низким энергопотреблением, малым размером кода и малой площадью кристалла.

3.1.2. Системы команд ARM и Thumb

Система команд ARM — это набор 32-битных инструкций, предоставляющий полную функциональность для обработки и управления данными.

Система команд Thumb изначально была разработана, как набор 16-битных инструкций, подробно описанный в [6], отражающий лишь подмножество функциональности системы команд ARM. Этот набор значительно увеличивает плотность кода в обмен за небольшую потерю в производительности. Процессор, исполняющий Thumb инструкции, может переключаться на исполнение системы команд ARM на лету для участков кода, критичных к показателю производительности, в частности для управления прерываниями. Начиная с версии ARMv6T2, архитектура поддерживает технологию Thumb-2. Она позволяет расширять оригинальные 16-битные Thumb-инструкции до 32-битных, что приводит к достижению производительности, близкой к ARM, с еще большей плотностью кода, по сравнению с предыдущей версией системы команд.

В последних версиях ARM архитектур функциональность системы команд ARM и Thumb практически не отличаются.

3.1.3. Разновидности системы команд ARMv7 и режимы работы

Системы команд ARM и Thumb претерпели значительные изменения с тех пор, как они впервые были разработаны, и они будут развиваться в будущем. Семь главных версий системы команд (ARMv1-ARMv7) были реализованы до сегодняшнего момента. Первые три из них уже значительно устарели и более не поддерживаются. Нас будет интересовать последняя версия ARMv7. Процессор этой версии имеет три разновидности режима исполнения команд:

ARMv7-A: Режим приложений. Реализован в виде традиционной системы команд ARM с большим количеством состояний. В этом режиме осуществлена поддержка VMSA (Virtual Memory System Architecture), основанная на MMU, а также вышеописанные системы команд ARM и Thumb.

ARMv7-R: Системный режим. Реализован в виде традиционной системы команд ARM с большим количеством состояний. В этом режиме осуществлена поддержка PMSA (Protected Memory System Architecture), основанная на MPU, а также вышеописанные системы команд ARM и Thumb.

ARMv7-M: Микроконтроллерный режим. Создан для быстрой обработки прерываний и реализован в виде модели с поддержкой аппаратного управления регистрами и управления прерываниями с помощью языков программирования высокого уровня. Реализован в виде ARMv7 PMSA и поддерживает только систему команд Thumb-2.

3.1.4. Расширения системы команд

У системы команд ARM и Thumb существуют следующие расширения:

ThumbEE. Это вариант системы команд Thumb, созданный в качестве цели для динамически сгенерированного кода. Расширение является:

- обязательным для режима ARMv7-A;
- опциональным для режима ARMv7-R.

VFP. Сопроцессорное расширение для работы с действительными числами. Существуют три версии этого расширения:

1. VFPv1 является устаревшим.
2. VFPv2 является опциональным расширением к:
 - системе команд ARM для архитектур ARMv5TE, ARMv5TEJ, ARMv6, ARMv6K;
 - системе команд ARM и Thumb для архитектуры ARMv6T2.
3. VFPv3 является опциональным расширением к системам команд ARM, Thumb

и ThumbEE в режимах ARMv7-A и ARMv7-R. VFPv3 может быть использовано как с 32-битными, так и с 16-битными регистрами, а также оно может быть расширено инструкциями, работающими с действительными числами половинной точности.

Усовершенствованный SIMD. Это набор команд, предоставляющий ОКМД (одиночный поток команд, множественный поток данных) функциональность. Является опциональным расширением к ARMv7-A и ARMv7-R режимам. В случае, если SIMD и VFPv3 включены одновременно, то они используют общий набор регистров и имеют некоторые общие инструкции. SIMD может быть также расширено инструкциями, работающими с действительными числами половинной точности.

Расширения безопасности. Это набор функций для облегчения разработки приложений, требующих высокого уровня защиты. Этот набор является опциональным расширением к архитектуре ARMv6K и режиму ARMv7-A.

Jazelle. Это набор инструкций для поддержки исполнения байт-кода Java, расширяющий ARMv5TE до ARMv5TEJ. Начиная с версии ARMv6, Jazelle является неотъемлемой частью архитектуры, но все равно описывается, как отдельное расширение.

Мультипроцессорное расширение. Это набор функций, разработанный для улучшения многопроцессорной функциональности. Он является опциональным расширением к режимам ARMv7-R и ARMv7-A.

3.1.5. Модель памяти архитектуры ARM

ARM архитектура использует единое плоское адресное пространство размером 232 байт. Также его можно рассматривать в качестве 230 32-битных слов или 231 16-битных полуслов.

Архитектура предоставляет средства для:

- предотвращения невыровненного доступа к памяти;
- ограничения доступа приложений к указанной области памяти;
- перевода виртуальных адресов, полученных после исполнения инструкций, в физические адреса;
- переключения интерпретации слова и полуслова между обратным и прямым порядком байтов;
- предотвращения неупорядоченного доступа к памяти;
- контроля кэша;
- синхронизации доступа к общей памяти несколькими процессорами.

3.2. Технология архитектурного профилирования

На начальных стадиях разработки проблемно-ориентированных процессоров, такие данные как: частота использования целочисленных операций и операций с плавающей запятой, динамическая битовая длина некоторых типов данных, поведение ветвления и доступа к памяти и т.д. — могут быть использованы для принятия важных решений, относящихся к битовой длине инструкций, размеру регистрового файла, структуре кэш памяти и т.д. Процесс сбора всей этой информации называется архитектурным профилированием и осуществляется с помощью программы, называемой профайлером. В данной работе для исследования эффективности системы команд ARM архитектуры автором была реализована технологическая цепочка, показанная на рис. 9.

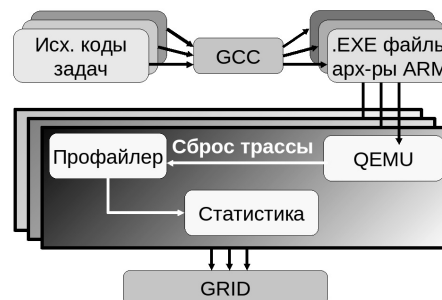


Рис. 9: Технологическая цепочка

Исходные коды SPEC тестов проходят процесс кросс-компиляции с хост-платформой x86 с помощью компилятора GCC, после чего получаются исполняемые файлы архитектуры ARM. Далее, эти исполняемые файлы запускаются на эмуляторе QEMU, который, в свою очередь, периодически сбрасывает трассу исполнения этих тестов в бинарном виде и через конвейер (FIFO файл) передает ее профайлеру, также реализованного автором. Эта программа декодирует полученные инструкции, анализирует их и увеличивает соответствующие счетчики. В результате работы профайлера получают разного рода статистики распределения частоты использования инструкций и их атрибутов. Для каждого теста эта цепочка оформляется в качестве отдельной независимой задачи и запускается для ускорения работы на GRID системе Condor.

3.3. Инструментарий

В этой главе рассмотрены подробно технологии, примененные автором для реализации вышеописанной цепочки, а также целесообразность их использования в данной работе.

3.3.1. Эмулятор аппаратного обеспечения различных платформ QEMU

QEMU — это программа с открытым исходным кодом для эмуляции аппаратного обеспечения различных платформ (в том числе и ARM). Она может запускать немодифицированную целевую операционную систему (например, Windows или Linux) и все ее приложения на виртуальной машине. Сам по себе QEMU может быть запущен на различных операционных системах, таких как Linux, Windows and Mac OS X. Целевой и хост-процессоры могут быть различными.

Главное назначение QEMU запускать одну операционную систему под другой, например, Linux под Windows и наоборот. Другое назначение QEMU — это процесс отладки, т. к. виртуальная машина может быть легко остановлена, а ее состояние сохранено для

дальнейшего изучения. Открытость исходного кода позволяет, добавлением машинного описания, эмулировать любые другие специфические устройства.

QEMU также включает в себя эмуляцию пользовательского режима Linux системы. Эта функция позволяет запускать любой Linux процесс целевой платформы под другой хост-платформой, что в основном используется для тестирования результата кросс-компиляции или для эмуляции работы непосредственно процессора без необходимости запуска виртуальной машины.

QEMU состоит из нескольких подсистем:

- Эмулятора платформ.
- Эмулируемых устройств.
- Семейства устройств для подключения эмулируемых устройств к соответствующим хост-устройствам.
- Отладчика.
- Пользовательского интерфейса.

QEMU использует динамическую трансляцию, преобразуя на лету инструкции целевой платформы в инструкции хост-платформы. Результирующий двоичный код сохраняется в кэше для дальнейшего переиспользования. Преимущество динамической трансляции перед интерпретацией в том, что целевые инструкции извлекаются и декодируются только единожды.

Обычно динамические трансляторы тяжело портируются с одной хост-платформы на другую, потому что весь генератор кода должен быть полностью переписан под новую архитектуру. Количество работы, затрачиваемое на эту процедуру, сравнимо с работой, необходимой для добавления новой целевой архитектуры C компилятору. QEMU устроен намного проще, потому что он лишь соединяет части машинного кода, сгенерированного отдельно компилятором GCC [7].

QEMU обладает достаточно высокой производительностью. Для измерения эффективности процесса эмуляции x86 кода под x86 хост-платформой использовались

ВУТEmark бенчмарки [8]. В пользовательском режиме эмуляция целочисленных операций оказалась в 4 раза медленнее нативного исполнения, а с плавающей запятой — в 10 раз. В режиме полной эмуляции системы QEMU примерно в 30 раз быстрее Bochs [9], а в пользовательском режиме в 1.2 раза быстрее, чем valgrind [10] с опцией `-skin=none`. Эта опция отключает генерацию дополнительного кода для отладки.

Как уже упоминалось в вышеописанной главе, для запуска SPEC тестов, скомпилированных под ARM архитектуру, был использован именно эмулятор QEMU. Самой главной причиной целесообразности использования эмулятора, является отсутствие возможности запуска этих тестов на ARM машине. Дело в том, что SPEC бенчмарки являются слишком ресурсоемкими и практически ни одна ARM машина не рассчитана на нагрузку такой величины.

Таким образом, QEMU был выбран в качестве эмулятора ARM архитектуры по следующим причинам:

- Отсутствие возможности запуска SPEC тестов на ARM машине.
- Свободное распространение.
- Открытость исходного кода.
- Высокая производительность.
- Поддержка большого количества целевых и хост-платформ.

3.3.2. Система распределенных вычислений Condor

Одним из существенных минусов использования эмулятора является значительное замедление исполнения кода. В среднем для полного завершения работы одного SPEC бенчмарка на эмуляторе требуется несколько дней, поэтому целесообразно распараллелить этот процесс используя GRID систему. В данной работе предлагается использовать систему распределенных вычислений Condor.

Condor — это специализированная система управления ресурсами для задач, требующих больших вычислительных затрат,

включающая в себя механизм управления процессами, стратегию планирования, систему приоритетов, а также механизмы мониторинга и распределения вычислительных ресурсов. Пользователи вставляют в очередь свои задачи, а Condor решает, когда и где запускать их, основываясь на стратегии планирования, а также следит за прогрессом этих задач и в конечном итоге сообщает пользователю об их завершении.

Помимо предоставления традиционной функциональности систем, в основе которых лежит очередь пакетной обработки данных, новая архитектура Condor и ее уникальные механизмы позволяют ей добиться высокой производительности в тех областях, где обычные системы управления ресурсами справляются слабо, а именно в сфере вычислений с высокой пропускной способностью и оппортунистических вычислений. Цель первого направления — предоставить большое количество отказоустойчивой вычислительной мощности в течении длительного промежутка времени, эффективно используя доступные сетевые ресурсы. Цель второго — предоставить возможность использования вычислительных ресурсов без требования 100% доступности к ним. Эти два направления тесно связаны между собой, т.к. вычислений с высокой пропускной способностью можно легко добиться с помощью идеологии оппортунистического компьютеринга.

Некоторые механизмы Condor включают в себя следующее:

1. **Классады:** Механизм классадов предоставляет очень ясный и гибкий интерфейс для поиска подходящих вычислительных машин, требуемых выполняемой задачей. Он позволяет адаптироваться Condor-системе к любой желаемой политике использования ресурсов и соответствующим образом менять стратегию планирования в процессе объединения GRID узлов в единую вычислительную сеть.
2. **Контрольная точка задачи и миграция:** Во время исполнения некото-

рых типов задач, Condor способен периодически записывать их текущее состояние в файл и при необходимости восстановить работу приложения с этой контрольной точки, повышая тем самым, отказоустойчивость системы. Также, эта функциональность позволяет перенести задачу с одной машины на другую.

3. **Удаленные системные вызовы:** В процессе выполнения задач на удаленных машинах, Condor умеет сохранять локальное окружение процесса с помощью удаленных системных вызовов. Этот механизм мобильной изолированной программной среды позволяет перенаправлять все системные вызовы ввода-вывода задачи обратно на машину, с которой был осуществлен ее запуск. Таким образом, пользователю не нужно создавать файлы на удаленных машинах перед запуском Condor системы даже в отсутствие общей файловой системы.

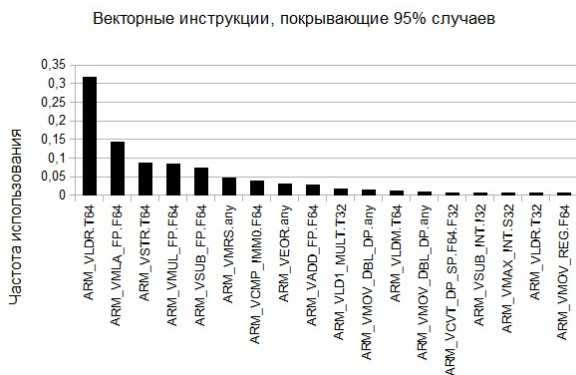


Рис. 10: Результаты анализа векторных инструкций

3.4. Результаты

Как уже упоминалось выше, у системы команд ARM архитектуры есть большое количество расширений. В частности, нас будут интересовать наборы векторных инструкций SIMD и VFP, так как они оказывают существенное влияние на производительность процессора. Результаты их анализа показали, что примерно из 1000 все-

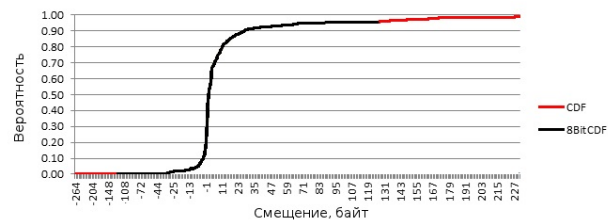


Рис. 11: Функция распределения частоты встречи LDR/STR инструкций в зависимости от величины относительного смещения в байтах

возможных комбинаций векторных инструкций и типов их атрибутов лишь 80 встретились в коде хотя бы раз, а 18 из них, как показано на рис. 10, покрывают 95% случаев использования. Также из рисунка видно, что VLDR/VSTR инструкции, выполняющие загрузку данных из памяти в регистр и наоборот соответственно, оказались самыми часто встречающимися, что и следовало ожидать, из-за принадлежности ARM архитектуры к семейству RISC. По этой же причине был вызван интерес исследования обычных LDR/STR инструкций, а точнее зависимости частоты их встречи в коде от количества относительного смещения в байтах, закодированного в поле прямого адреса. Из рис. 11 видно, что в 95% случаев в поле прямого адреса закодировано число в диапазоне от -127 до 127. Это означает, что в большинстве используется всего лишь 8 бит этого поля (1 бит на знак) кодировки из 12-ти доступных, то есть, по сути дела, 4 бита лишние.

4. Выводы

В этой работе были освещены две актуальные проблемы, возникающие в процессе дизайна микроархитектур. Первая проблема заключается в том, что большое количество исходного кода SPEC тестов делают процесс их отладки очень сложным и долгим. Для решения этой проблемы автором было предложено использование новой технологии InterAspect, позволяющей значительно облегчить этот процесс. Суть второй проблемы — необходимость постоянного мониторинга за производительностью микроархитектуры и поиска возможностей для ее улучшения. С этой целью автором

было проведено исследование эффективности системы команд архитектуры ARM с помощью метода архитектурного профилирования. Результаты анализа показали узкие места в системе команд ARM архитектуры, которые уже были учтены в процессе дизайна платформы SMI.

Литература

- [1] Standard Performance Evaluation Corporation, официальный сайт. www.spec.org.
- [2] Официальный сайт InterAspect <http://www.fsl.cs.stonybrook.edu/interaspect/>.
- [3] Kiczales G., Hilsdale E., Hugunin J., Kersten M., Palm J. and Griswold W. G. An overview of AspectJ. In Proceedings of the 15th European Conference on Object-Oriented Programming (2001), LNCS, Vol. 2072, pp. 327-355.
- [4] AT&T Research Labs. Graphviz, 2009. <http://www.graphviz.org>.
- [5] Huang X., Seyster J., Callanan S., Dixit K., Grosu R., Smolka S.A., Stoller S.D. and Zadok E. Software monitoring with controllable overhead. International Journal on Software Tools for Technology Transfer (STTT) (2010). Accepted for publication.
- [6] Gouge L., Segars S. Thumb: Reducing the Cost of 32-bit RISC Performance in Portable and Consumer Applications. Proceedings of the COMPCON Spring'96 — 41st IEEE International Computer Conference, 1996.
- [7] GNU Compiler Collection. <http://gcc.gnu.org/>.
- [8] The BYTEmark benchmark program, BYTE Magazine, Linux-версия доступна на <http://www.tux.org/mayer/linux/bmark.html>.
- [9] Kevin Lawton et al., the Bochs IA-32 Emulator Project, <http://bochs.sourceforge.net>.
- [10] Julian Seward et al., Valgrind, an open-source memory debugger for x86-GNU/Linux, <http://valgrind.kde.org/>.

Исследование Ni_3Hf как легирующего компонента для создания высокотемпературного сплава с эффектом памяти формы на основе Ni_3Ta *

Н. Н. Злобин
nikolay@zlobin.biz

Московский физико-технический институт (государственный университет)

Долгое время неупругую деформацию считали полностью необратимой. В начале 60-х годов XX в. был открыт обширный класс металлических материалов, у которых элементарный акт неупругой деформации осуществляется за счет структурного превращения. Такие материалы обладают обратимостью неупругой деформации. Явление самопроизвольного восстановления формы — эффект памяти формы (ЭПФ) — может наблюдаться как в изотермических условиях, так и при температурных изменениях. При теплосменах такие металлические материалы могут многократно обратимо деформироваться.

Ключевые слова: легирующий компонент, высокотемпературный сплав, эффект памяти формы.

Investigation of Ni_3Hf as an alloying component for the creation of a high-temperature Ni_3Ta -based memory alloy*

N. N. Zlobin

Moscow Institute of Physics and Technology (State University)

For a long time the inelastic deformation was considered completely irreversible. In the early 60-ies of the XX century. An extensive class of metallic materials was discovered, in which the elementary act of inelastic deformation is carried out through structural transformation. Such materials exhibit invertibility of inelastic deformation. The phenomenon of spontaneous shape recovery — shape memory effect (ESR) — can be observed both under isothermal conditions and under temperature changes. With thermal changes such metallic materials can be reversibly deformed many times.

Key words: alloying component, high-temperature alloy, form memory effect.

Введение

Мартенситное превращение, полиморфное превращение, при котором изменение взаимного расположения составляющих кристалл атомов (или молекул) происходит путём их упорядоченного перемещения, причем относительные смещения соседних атомов малы по сравнению с межаутомным расстоянием. Перестройка кристаллической решётки в микрообластях обычно сводится к деформации её ячейки, и конечная фаза мартенситного превращения может рассматриваться как однородно деформированная исходная фаза. Величина деформации мала (порядка 1-10%) и соответственно мал, по сравнению с энергией связи в кристалле, энергетический барьер, препятствующий однородному переходу исходной фазы в конечную. Необходимое условие мартенситного превращения, которое развивается путём образования и

роста областей более стабильной фазы в метастабильной, — сохранение упорядоченного контакта между фазами. Упорядоченное строение межфазных границ при малости барьера для однородного фазового перехода обеспечивает их малую энергию и высокую подвижность. Как следствие, избыточная энергия, необходимая для зарождения кристаллов новой фазы (мартенситных кристаллов), мала и при некотором отклонении от равновесия фаз становится сопоставимой с энергией дефектов, присутствующих в исходной фазе. Поэтому зарождение мартенситных кристаллов происходит с большой скоростью и может не требовать тепловых флуктуаций. Вследствие воздействия образовавшейся фазы на исходную фазу энергетический барьер для перемещения границы фаз существенно меньше, чем для однородного перехода; при небольших отклонениях от равновесия он исчезает — кристалл растёт со скоростью порядка звуковой и без тепловой активации (превраще-

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 16-07-00089, 18-07-20030.

ние возможно при температурах, близких к абсолютному нулю).

Мартенситные превращения обнаружены во многих кристаллических материалах: чистых металлах, многочисленных сплавах, ионных, ковалентных и молекулярных кристаллах. Наиболее полно изучены мартенситные превращения в сплавах на основе железа, в частности в связи с закалкой стали. Большие перспективы практического применения имеют возможность большого обратимого формоизменения при мартенситных превращениях (например, создание «сверхупругих» сплавов и изделий, восстанавливающих первоначальную форму при нагреве после пластической деформации — «эффект памяти»), а также связь мартенситных превращений с появлением сверхпроводящих свойств в некоторых металлах. Мартенситные превращения (часто в сочетании с диффузионным перераспределением компонентов и изменением атомного порядка) составляют основу многочисленных структурных превращений, благодаря которым с помощью термической и механической обработки осуществляется направленное изменение свойств кристаллических материалов. Значительный вклад в изучение мартенситных превращений внесли работы советских учёных (Г. В. Курдюмов и его школа).

1. Мартенситное превращение. Общие характеристики

Известно, что при закалке стали от высокой температуры (из аустенитной области) происходит повышение её твердости. После этого под микроскопом можно наблюдать специфическую структуру, которую называют мартенситом по имени немецкого исследователя Мартенса, впервые наблюдавшего её [1, 2].

Наиболее точно и просто мартенситное превращение можно определить как превращение решетки посредством деформации сдвига на основе кооперативного движения атомов (в исходной фазе атомы движутся не по отдельности, а как единый ком-

плекс, т. е. деформация сдвига осуществляется как бы «ряд за рядом»). В результате образуются линзообразные или пластинчатые области [1, 2].

1.1. Особенности мартенситного превращения [1, 2]

1. Мартенситная фаза является однородным твердым раствором замещения или внедрения.
2. Превращение является бездиффузионным. Концентрация растворенных атомов, находящихся в твердом в растворе мартенситной фазе такая же, как и в исходной фазе, и диффузии атомов на большие расстояния не происходит.
3. Превращение характеризуется деформацией формы (или поверхностным рельефом), при этом она имеет постоянную величину.
4. Кристаллы мартенсита имеют определённую плоскость габитуса (плоскость габитуса определяют индексами кристаллографических плоскостей исходной фазы)
5. Между решетками исходной и мартенситной фаз существуют определенные ориентационные соотношения.
6. В решетке мартенситной фазы всегда существуют дефекты. Для объяснения дефектов вводят дополнительную деформацию (её называют деформацией с инвариантной решеткой): скольжение или двойникование (рис. 1). Обнаруживают её следы (дислокации, дефекты упаковки), обычно, с помощью электронного микроскопа.

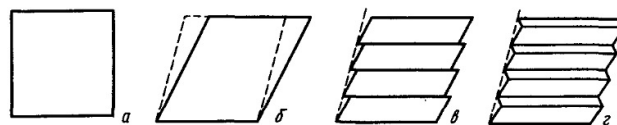


Рис. 1: Деформация решетки и дополнительная деформация скольжением (в) или двойникованием (г). Штриховой линией показано действительное изменение формы

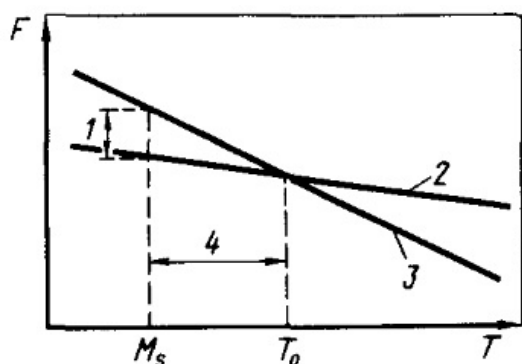


Рис. 2: Зависимость химической свободной энергии F исходной и мартенситной фаз от температуры и условия мартенситного превращения: 1 – движущая сила (энергия поверхности раздела, энергия деформации и т. п.); 2 – мартенсит (М); 3 – исходная фаза (Р); 4 – степень переохлаждения. Характерные значения степени переохлаждения достигают $\sim 200^\circ\text{C}$ (например, сплавы с Fe) и $\sim 5\text{--}30^\circ\text{C}$ для сплавов с ЭПФ

1.2. Принципы классификации превращений [1, 2]

- Термоупругий или нетермоупругий характер превращения.
- Кинетические признаки: скорость превращения (адиабатический и изотермический МП), взрывной или плавный его характер; обратимость (обратимые и необратимые МП), величина гистерезиса между прямым и обратным превращениями.
- Род превращения (превращения первого рода; превращения, близкие к превращениям второго рода).
- Химический состав материала (превращения в чистых элементах, превращения в сплавах, в соединениях различного типа — ионных солях, ковалентных соединениях).
- Термодинамические условия превращения (при изменении температуры, давления, магнитного или электрического поля).
- Превращения в статических и динамических внешних полях.
- Кристаллографические характеристики (структура исходной и мартенситной фаз).
- Размерные характеристики объектов (превращения в монокристаллах, поли-

кристаллах, в пленках, в микро- и нанокристаллических материалах).

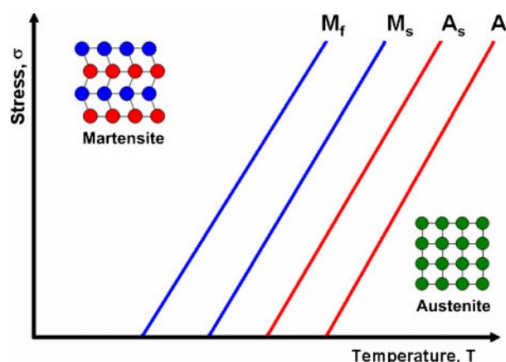


Рис. 3: Иллюстрация линейной зависимости между приложенным напряжением и температурой превращения. При температуре выше A_f образец полностью находится в аустенитной фазе, а при температуре ниже M_f — полностью в мартенситной фазе

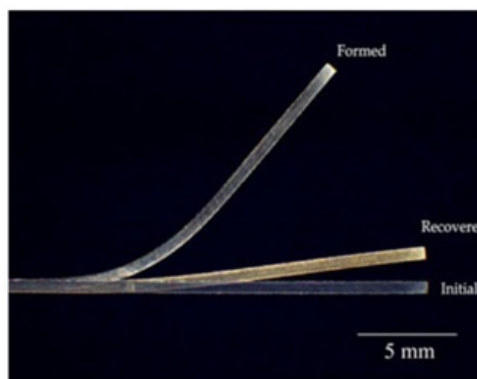


Рис. 4: Иллюстрация одностороннего эффекта памяти формы: данный образец (initial), находящийся в мартенситной фазе, деформируют (formed), но после нагревания он в состоянии вернуться в практически исходную форму (recovered)

1.3. Термодинамика МП [1, 2]

Для превращения исходной фазы в мартенситную необходима так называемая движущая сила, то есть, чтобы свободная энергия нехимической природы (энергия деформации, энергия поверхности раздела) была больше, чем разница химических энергий исходной и мартенситной фаз.

Обычно вводят следующие обозначения:

- M_s – температура начала прямого превращения;
- M_f – температура окончания прямого превращения;

- T_0 – температура равновесия (химические свободные энергии исходной и мартенситной фаз равны);
- A_s – температура начала обратного превращения;
- A_f – температура окончания обратного превращения.

Прямое превращение не происходит, если нет переохлаждения до температуры M_s , меньшей, чем T_0 (рис. 2). Таким же образом движущая сила необходима и для обратного превращения, т. е. необходим перегрев до температуры A_s , большей, чем T_0 . Пример линейной зависимости между приложенным напряжением и температурой превращения представлен на рис. 3.

2. Эффект памяти формы

2.1. Общие характеристики

При деформировании образца в мартенситной фазе, т. е. при $T < A_f$, даже при снятии нагрузки деформация полностью не устраняется. Однако, остаточная деформация исчезает почти полностью при нагреве образца выше температуры A_f . Под этим процессом, согласно определению, и понимают эффект памяти формы (рис. 4) [3, 4]. ЭПФ наблюдается в полной форме (т. е. восстанавливается как кристаллическая структура, так и кристаллографическая ориентировка исходной фазы перед превращением) только в сплавах с термоупругим характером превращений.

Для полного восстановления необходимы:

- кристаллографическая обратимость.
- процесс деформации должен осуществляться без участия скольжения (т. к. это необратимый процесс).

2.2. Механизм эффекта памяти формы

Основным механизмом деформации сплавов с ЭПФ находящихся полностью в мартенситном состоянии, независимо от типа внутренних дефектов является деформация двойникованием. Деформация, которая может быть восстановлена в результате

действия эффекта памяти формы обусловлена взаимодействием между двойниковыми доменами. Деформация развивается до тех пор, пока весь образец в целом не становится состоящим из одного домена [3, 4].

2.3. Двухсторонний эффект памяти формы

Следует отметить, что в рассмотренном выше ЭПФ, который имеет название одностороннего, запоминается форма образца только в состоянии исходной фазы.

Однако, наряду с однократным ЭПФ существует обратимый (двухсторонний) ЭПФ (ОЭПФ), заключающийся в обратимом (т. е. «запоминается» состояние обеих фаз) изменении формы при термоциклировании (нагреве выше температуры A_f и охлаждении ниже температуры M_f) через интервал мартенситных превращений (рис. 5). Источником ОЭПФ служат поля напряжений от дислокационной субструктуры и когерентных выделений, которые оказывают ориентирующее действие на мартенситное превращение [3, 4].

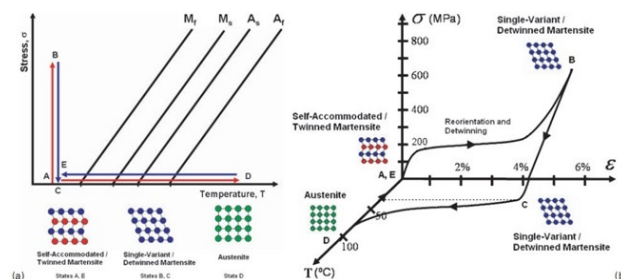


Рис. 5: Иллюстрация эффекта памяти формы при помощи зависимости «напряжение-температура» (а) и «напряжение-деформация» (б)

Образец, первоначально находящийся полностью в мартенситной фазе (состояние А), деформируют при температуре ниже M_f , что приводит к изменению его внутренней структуры (состояние В) и сохранению этой структуры после снятия внешнего напряжения (состояние С). Это приводит к внешнему изменению формы (показано на рис. 5 б). После нагрева выше температуры A_f мартенсит полностью превращается в аустенит (состояние D), который, в

свою очередь, снова превращается в мартенсит при охлаждении ниже температуры M_f , восстанавливая первоначальную форму.

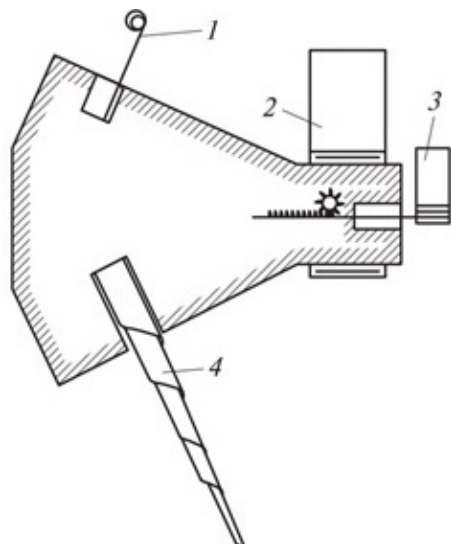


Рис. 6: Схема космического аппарата с саморазворачивающимися элементами:

- 1 — антенна;
- 2 — солнечная батарея;
- 3 — излучатель энергии;
- 3 — механический стабилизатор;
- 4 — механический стабилизатор

3. Применение материалов, которые обладают МП и демонстрируют ЭПФ

3.1. Применение в технике

Материалы с эффектом памяти формы известны уже более 50 лет, однако, интерес к ним не ослабевает и в настоящее время. Например, в космической технике с помощью этих сплавов эффективно решается традиционная проблема экономии места (рис. 6). Свернутые в компактную форму антенны, механизмы стабилизации, солнечные батареи распрямляются или выдвигаются из отсеков корпуса после запуска аппарата на орбиту от действия солнечного тепла. Совмещая силовые и деформационные свойства элементов из металла с эффектом памяти формы, удастся проектировать исключительно простые и эффективные исполнительные устройства роботов, разнообразные приводы в конвейерных

производствах, усилители перемещений и т. д. [4, 5].



Рис. 7: Термомеханические соединения (ТМС)

Еще одна область применения касается создания плотных и неразъемных соединений: можно скреплять трубы и стержни надетыми на них муфтами из сплавов с памятью формы (рис. 7). Здесь используется необычная с инженерной точки зрения особенность проявления памяти металла, когда мы заставляем его сжиматься при нагреве. Таким же образом можно производить опрессовку деталей, используя память формы как свойство инструмента; собирать узлы, не поддающиеся этой операции с помощью традиционной технологии. Например, трубы диаметром 20 мм легко скрепляются наружной (стягивающей) или внутренней распорной муфтой из никелида титана при ее толщине около 2 мм. Трубы при этом выдерживают внутреннее давление 200 атм и более. Подобным способом удастся скреплять металл с резиной, пластмассой или керамикой. Подобные термомеханические соединения (ТМС) были использованы при сборке 14,5 м ферменной конструкции (массой – 90 кг) астрофизического модуля «Квант» станции «Мир». Также широко применяются термочувствительные элементы в клапанах различного назначения, которые например, предотвращают замерзание воды в трубах или взрыв газа при пожаре.

Ещё один яркий пример успешного применения сплавов с эффектом памяти формы — компания Боинг, использовавшая их при

проектировании различных областей реактивного сопла в виду преимуществ сплавов с ЭПФ при использовании специфического топлива, а также при разработке конструкций для снижения шума от турбин при взлёте. Так, для достижения шумоподавления используют так называемые шевроны (англ. The variable geometry Chevrons) — зубчатые аэродинамические устройства, располагающиеся вдоль задней кромки реактивной турбины (рис. 8) [6].



Рис. 8: Компания Boeing использует приводы на основе сплавов с ЭПФ для того, чтобы видоизменять форму шевронов для удовлетворения требованиям акустических тестов по снижению шума при взлете самолетов. Основная идея заключается в смешивании потока воздуха из сопла, окружающего и первоначального (захватываемого в сопло) потоков:

а) — иллюстрирует желаемые конфигурации выхлопного сопла с шевронами;

б) — изогнутые приводы на основе сплавов с ЭПФ контролируют отклонение шевронов;

в) — тестирование турбин с шевронами в реальных условиях перед началом полета

3.2. Применение в медицине

Сегодня о стендах для расширения просвета кровеносных сосудов слышали практически все. Аналогичным способом проводят расширение и восстанавливают проходимость полых органов: желчных протоков, пищевода, трахеи и др. Для каждого случая хирургического вмешательства были разработаны оригинальные конструкции протезов с ЭПФ и сверхупругостью из сплавов Ti-Ni (см. рис. 9 а) и доставляющих устройств (на рисунке не показаны). Эти разработки также получили приоритет в мире [5].

Альтернативная задача — обжатие вены — также решается устройством с ЭПФ. Экстравазальный корректор (ЭВК) в удобной для установки форме подводят к поражен-

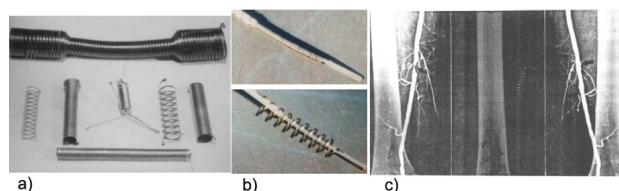


Рис. 9: а) — типы эндопротезов; б) — протез в компактном виде и после нагрева до 37°С; в) — протезирование бедренной артерии, стрелками показаны место сужения (слева) и восстановленный кровоток (справа)

ному участку вены. При контакте с веной корректор под действием тепла крови (тела) пациента «вспоминает» свою исходно заданную форму и параметры, обхватывая вену и самофиксируясь на ней. Его применение позволяет восстанавливать функцию несостоятельных клапанов глубоких вен нижних конечностей, предотвращая опасность обратного оттока и застоя крови (варикозная болезнь). Устройство устанавливается интраоперационно с минимальной травмой для пациента.

Сочетание ЭПФ и обратимого ЭПФ использовано при создании зажима «Клест» для клипирования сосудов и мягких тканей при лапароскопических операциях. Клипса «Клест» предназначена для экстренной и плановой остановки внутренних кровотечений. Она является самосрабатывающей, обеспечивает 100% — ную надежность перекрытия сосудов при лапароскопическом и открытом применении и постояннодействующую компрессию, при ее установке нет необходимости в применении дорогостоящего сложного манипулятора. При необходимости клипса может быть легко удалена за счет реализации специально наведенного обратимого эффекта памяти формы (ОЭПФ) (рис. 10).

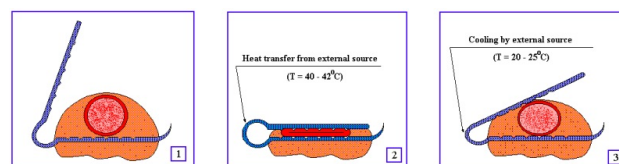


Рис. 10: Схема действия клипсы: 1 — охват сосуда с прошиванием тканей, 2 — обжатие сосуда за счет ЭПФ, 3 — освобождение сосуда за счет ОЭПФ

Свойство сверхупругости было использовано при создании ловушки «Трал» для извлечения камней из полых органов и скрепок для сшивания кровеносных сосудов при аортокоронарном шунтировании. «Трал» в отличие от существующих аналогов обеспечивает легкое освобождение от камней простым выпрямлением спиральной бранши. Затем коническая форма ловушки восстанавливается за счет сверхупругости (рис. 11 а).

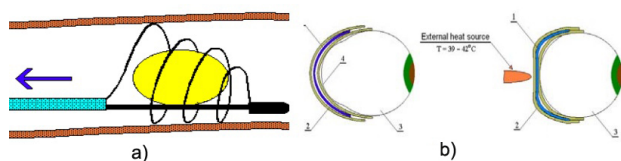


Рис. 11: а) – схема действия сверхупругого экстрактора «Трал»; б) – схема работы динамической пломбы

Перспективной разработкой для офтальмологии, использующей свойства памяти и сверхупругости, является динамическая пломба для лечения высокой осложненной близорукости путем восстановления формы глазного яблока (рис. 11 б). При охлаждении устройство в виде пластины (1) или пружины принудительно деформируют в плоскую форму и в таком состоянии размещают и фиксируют между лентами трансплантата (2). Ленточный трансплантат размещают на задней стенке глазного яблока (3) в области патологической выпуклости и подшивают к склере. При подведении к «динамической пломбе» (1) внешнего источника тепла устройство «вспоминает» исходно заданную форму и компенсирует патологическую выпуклость глазного яблока.

4. Высокотемпературные сплавы с эффектом памяти формы

Принятое нами определение понятия «высокотемпературного сплава с эффектом памяти формы» (см. раздел определения) может включать сплавы, которые демонстрируют прямое МП при температурах ниже 390К, но исключает сплавы, в которых происходит обратное МП при температурах

выше 390К только при первом нагревательном цикле из-за механической или термической стабилизации или такие, которые проявляют псевдо-упругий характер МП выше 390К [7].

Высокотемпературные сплавы с эффектом памяти формы (ВСЭПФ) могут упростить и улучшить операционную эффективность многих механических компонентов и устройств, созданных для работы при температурах выше 100°C, потенциально — в автомобильной, аэрокосмической, производственной и других энергоемких отраслях.

Помимо высоких температур перехода, потенциальные высокотемпературные сплавы с эффектом памяти формы также должны проявлять приемлемый уровень восстанавливаемой деформации, долгосрочную устойчивость, сопротивляемость к пластической деформации и деформации скольжением и подходящую сопротивляемость другим факторам окружающей среды. Удовлетворять этим критериям станет все сложнее и сложнее в виду постоянного роста диапазона температур, при которых работают механизмы, нуждающиеся в ВСЭПФ. Более того, слабая работоспособность многих ВСЭПФ, обуславливаемая высокой упорядоченностью этих интерметаллических структур, а также высокая себестоимость материалов создают дополнительные проблемы для их коммерческого использования.

4.1. Особенности сплавов с ЭПФ при высоких температурах

Помимо факторов, характерных для всех сплавов испытывающих МП и проявляющих ЭПФ, существует несколько других, которые являются существенными для рассмотрения именно при высоких температурах [7].

4.1.1 Механические характеристики при высоких температурах

4.1.1.а. Влияние температуры на предел текучести (напряжение текуче-

сти)

Деформационные характеристики ВСЭПФ усложняются за счет доступности термической энергии при высоких температурах. Общая проблема всех ВСЭПФ заключается в том, чтобы минимизировать величину невосстанавливаемой деформации, возникающей из-за пластической деформации, которая присутствует при фазовом превращении. Эта проблема обостряется при высоких температурах по причине уменьшения модуля напряжения как для аустенитной, так и для мартенситной фазы.

В результате, ВСЭПФ очень редко проявляют заметную память формы или сверхупругость, а в тех сплавах, которые даже это делают, полное восстановление при любой величине приложенного в начале напряжения наблюдается крайне редко.

4.1.1.b. Деформация ползучести

Деформация ползучести является важной для ВСЭПФ с высокими температурами перехода (400°C). Хотя деформация ползучести и является нежизнеспособным механизмом при работе компонентов на основе ВСЭПФ с более низкими температурами перехода, её эффекты стоит учитывать при высокотемпературной обработке и образовании веществ.

Фазовое превращение, как правило, создает локальные напряжения многим большие, чем макроскопические и, как предполагают, сопротивление за счет ползучести окажется куда хуже в случае, когда сплав с ЭПФ испытывает циклы превращений при высоких температурах, по сравнению с сопротивлением статической деформации ползучести. Если для работы компонентов на основе ВСЭПФ требуются высокие величины приложенного напряжения или давления, деформация ползучести окажется проблемой даже при «средних» температурах (500°C). Необходимо учитывать, какое влияние оказывает фазовое превращение на деформацию ползучести также, как и то, как влияет эта деформация на само превращение. К сожалению, на данный

момент этот вопрос практически не исследован.

4.1.2. Микроструктурная неустойчивость

Фаза, возникающая в процессе превращения многих сплавов с ЭПФ является неравновесной, поэтому свойства ЭПФ многих сплавов зависят от возможности существования этой неравновесной фазы. Так, стремление к образованию равновесной фазы при высоких температурах может быть разрушительным для ЭПФ или сверхупругости. Равновесные фазы, которые образуются с течением времени, в большинстве случаев не способны испытывать превращение и обычно либо оказываются хрупкими, что приводит к потере прочности, либо очень пластичными, что негативно влияет на способность полного восстановления.

К сожалению, для многих ВСЭПФ процесс образования равновесных фаз может не быть контролируемым в достаточной степени. Если диапазон температур достаточно высок для осаждения таких равновесных фаз в достаточном количестве, то их образование приведет к непрерывному изменению свойств ЭПФ взятого сплава, обусловленному изменением строения кристаллической решетки. В конце концов, эти изменения достигнут такой степени, что ВСЭПФ больше не сможет выполнять предназначенную функцию.

4.1.3. Окисление

Сплавы, имеющие контакт с атмосферой при высоких температурах будут неизбежно взаимодействовать с кислородом и другими газами. В результате, на поверхности образца могут образовываться продукты окисления металла. Если реакция не будет остановлена и достаточное количество кислорода будет доступно, что имеет место в неконтролируемой атмосфере, все большая и большая масса сплава будет превращена в оксид. К счастью, оксидные формирования на поверхности замедляют диффузию металлических ионов и кислорода, а также «контролируют» уровень окисле-

ния. Следовательно, помимо температуры, уровень окисления определяется проницаемостью оксидного слоя. Сплавы, созданные для работы при высоких температурах, способны образовывать стойкие и непроницаемые оксидные слои, которые препятствуют дальнейшему окислению, начиная с момента, когда образовалась некоторая критическая толщина оксидного слоя.

Тем не менее, часто оксидный слой состоит по большинству из оксида лишь одного компонента сплава, тем самым уменьшая концентрацию этого компонента в кристалле и существенно понижая локальные температуры перехода.

Резюмируя сказанное выше, стоит сказать, что окисление не является проблемой при температурах перехода около 200-300°C, но является проблемой при более высоких температурах. Также стоит отметить, что уникальным фактором для сплавов с ЭПФ является чувствительность температур перехода к изменениям состава, связанным с окислением, которая ограничивает как максимальную температуру, так и время действия (экспозиции) компонентов с ЭПФ. Изменения формы, связанные с изменением кристалла, также могут вызвать разрушение хрупкого оксидного слоя, усложняя тем самым создание устойчивого слоя даже в случае, когда сплавы имеют высокую сопротивляемость окислению.

4.2. Применение ВСЭПФ: возможности и проблемы

Современно практическое применение сплавов с ЭПФ ограничивается температурами ниже 100°C. Это ограничение для превращения двух коммерчески наиболее успешных сплавов с ЭПФ: Ni-Ti и трехкомпонентные системы на основе Cu. В процессе обработки, требуемой для создания стабильных характеристик по памяти формы или сверхупругости, температуры превращения существенно падают.

Тем не менее, есть потребность в ВСЭПФ, работающих при температурах выше 390К,

потому что нужны материалы для работы систем по превращению энергии, компактные исполнительные механизмы, которые могут совершать существенную работу, при этом занимая малый объем. ВСЭПФ требуются также для роботизированной, автомобилестроительной отраслей, для использования в микроэлектромеханических системах (МЭМС: микромашины, кардиостимуляторы и дефибрилляторы), а также для раскрывающихся-закрывающихся компонентов в аэрокосмических устройствах.

Однако, в виду недостаточной устойчивости, пластичности, надежности и других функциональных характеристик до сих пор нет ни единого примера успешного применения ВСЭПФ.

К сожалению, существует не слишком много возможностей улучшить ситуацию. По сути, лишь легирование дополнительными компонентами для укрепления сплава и предотвращения пластической деформации выглядит в достаточной степени подходящим методом. Маловероятно, что отвердевающие растворы смогут оказаться решением, потому что они всегда связаны с усилением связей, которые делают высокотемпературную фазу более устойчивой или, другими словами, уменьшает температуры МП (Shabalovskaya, 1993; Firstov et al., 2003). Также, стоит отметить, что легирующие добавки, которые могут увеличить прочность, всегда ведут к повышению модуля упругости, что в свою очередь мешает высокотемпературному эффекту памяти формы (Firstov et al., 2004b).

5. Экспериментальные исследования и обсуждение результатов для Ni₃Ta и Ni₃Hf

5.1. Один из объектов исследования — Ni₃Ta. Свежие результаты

Согласно работе [8], было обнаружено мартенситное превращение, протекающее в интерметаллическом соединении Ni₃Ta. Ос-

Таблица 1:

	Нагрев					
	Пик №1		Пик №2		Пик №3	
	Температура пика, С	Кол-во теплоты, Дж/г	Температура пика, С	Кол-во теплоты, Дж/г	Температура пика, С	Кол-во теплоты, Дж/г
Ход №1	1196,31	20,13	1245,30	9,15	1269,54	1,28
Ход №2	1195,95	25,83	1244,17	3,023	1270,99	1,60
Ход №3	1196,07	25,87	1242,44	1,227	1271,20	1,40

новными аргументами в пользу этого вывода были:

- Подobie диаграмм фазовых состояний и принадлежность мартенситной структуры к той же пространственной группе, что и никелид титана (самый известный среди сплавов, испытывающих МП).
- Морфология кристаллов: оптическая металлография показала, что при комнатной температуре зёрна размером 20-100 мкм заполнены кристаллами копьевидной и пластинчатой форм (рис. 12).
- Стабильность характеристических температур перехода при многократном термоциклировании через интервал МП.

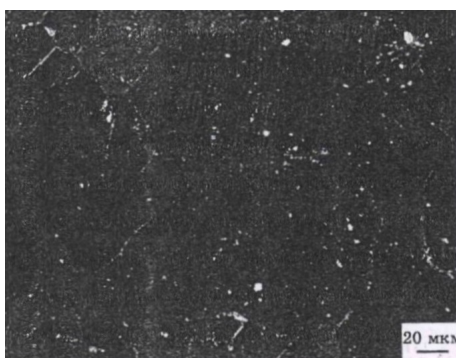


Рис. 12: Микроструктура слитка интерметаллида Ni₃Ta после вакуумного дугового переплава (оптическая металлография; поляризованный свет)

Низкая симметрия мартенситной фазы, высокая стабильность характеристик мартенситного превращения позволили надеяться на возможность реализации стабильного высокотемпературного эффекта памяти формы в этом соединении. Основными результатами более поздних работ было определение величины температурного гистерезиса (~50К) и характеристических

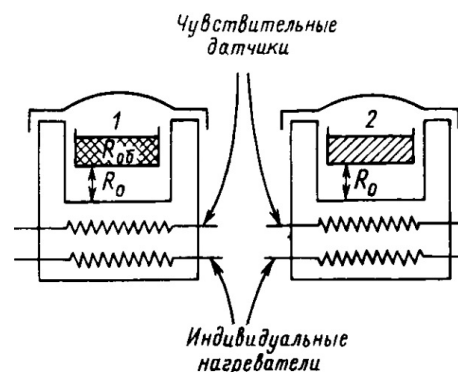


Рис. 13: Схематическое изображение калориметрических камер в ДСК: 1 – образец, 2 – эталон. Для исследования образца Ni₃Nf по методу ДСК использовался прибор SDT Q600 компании TA Instruments, год выпуска – 2008. Скорость нагревания, которая была использована в данной работе: 20°С/мин (рис. 14)



Рис. 14: Схематический вид установки, на которой проводились измерения по методу дифференциальной сканирующей калориметрии

температур перехода ($M_s=530\text{K}$, $M_f=500\text{K}$, $A_s=550\text{K}$, $A_f=580\text{K}$), 10% предварительной деформации за счет сжатия при комнатной температуре были практически полностью восстановлены после нагреванием с перегревом в 50К выше температуры A_s .

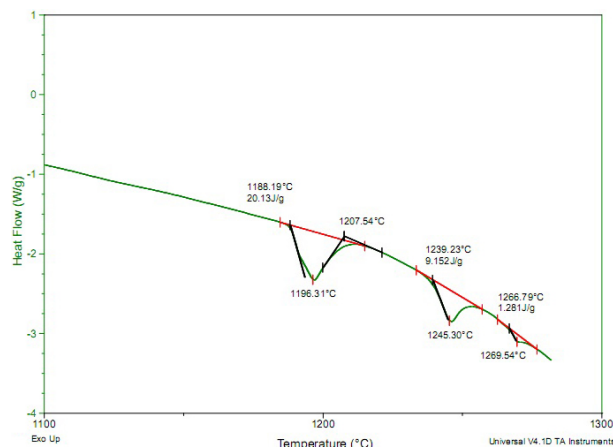


Рис. 15: Тепловой поток в зависимости от температуры, Ni_3Hf (нагрев №1)

5.2. Исследование Ni_3Hf

5.2.1. Метод дифференциальной сканирующей калориметрии

Две идентичные ячейки, изолированные от окружающей среды с помощью адиабатической оболочки, прогреваются с постоянной скоростью. При этом в одной из ячеек находится образец, а в другой - эталон. Мощности электронагревателей обеих ячеек автоматически регулируются так, чтобы выдержать заданную скорость нагрева, поэтому для поддержания одинаковой температуры в ячейках и сохранения одинаковой скорости их прогрева, мощности, подводимые к ячейкам различны (рис. 13). Экспериментальные кривые представляют собой зависимость теплового потока от температуры.

5.2.2. Сбор экспериментальных данных по методу дифференциальной сканирующей калориметрии для Ni_3Hf и их обсуждение

Было проведено по 3 цикла нагревания и охлаждения. Для каждого из циклов были выделены 3 пика, для которых при помощи соответствующего программного обеспечения определены основные параметры: температура пика и количество теплоты, соответствующее данному пику (рис. 15).

Данные были обработаны при помощи соответствующего программного обеспечения

[Universal TA Instruments] и систематизированы в таблицу 1.

Главный вывод: существует пик, который соответствует полиморфному превращению, а с нашей точки зрения — мартенситному. Таковым является Пик №1.

Основные аргументы:

- стабильность температур и количества теплоты, соответствующих данному пику при термоциклировании;
- температуры превращения соответствуют некому полиморфному превращению (данные анализа фазовой диаграммы из работы [9]);
- малый гистерезис (невязка между прямым и обратным ходом), равный $\sim 12-13^\circ\text{C}$.

Пики №2 и №3 имеют выраженные признаки диффузионного характера превращений.

Основные аргументы:

- нестабильность с точки зрения количества теплоты при превращениях им соответствующих;
- есть реакции, соответствующие данным пикам (данные проведенного анализа фазовых диаграмм из работы [10]).

5.2.3. Сбор данных по методу сканирующей электронной микроскопии (отраженные электроны) по Ni_3Hf

Ускоряющее напряжение прибора было равным 25 кВ. В качестве критериев для будущих выводов была учтена зависимость цвета участка на снимке не только от атомного номера элемента, а и от кристаллографической ориентации. В виде результатов были получены следующие снимки (рис. 16 и 17):

Основные аргументы в пользу наличия мартенситной фазы у Ni_3Hf :

- виден реечный контраст (подобный контраст замечался ранее на Ni_3Ta);
- трещины — косвенный аргумент того, что произошел сдвиг (необходимое условие МП);

- состав кошьевидной фазы 75-25(Ni-Hf), что позволяет, в виду схожести фазовых диаграмм в области данной стехиометрии, допустить аналогичные свойства с Ni₃Ta (данные о стехиометрии получены при помощи рентгеноструктурного микроанализа научным руководителем).

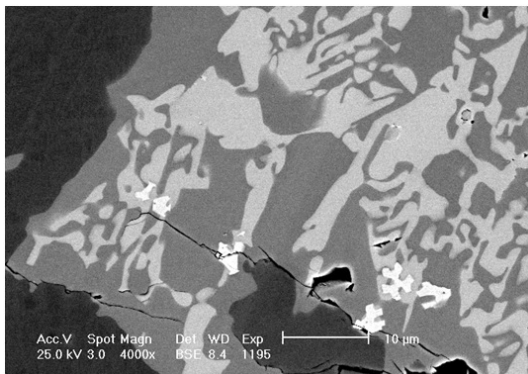


Рис. 16: Фотография образца Ni₃Hf в отраженных электронах, масштаб = 10 мкм

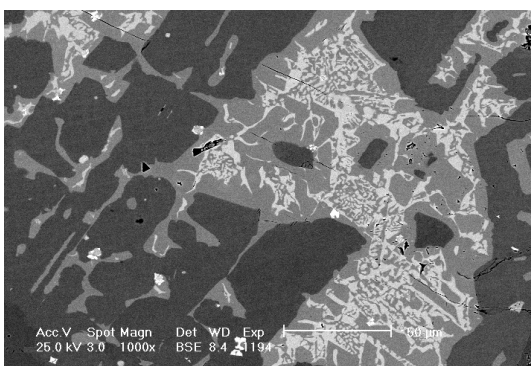


Рис. 17: Фотография образца Ni₃Hf в отраженных электронах, масштаб = 50 мкм

5.2.4. Проверка памяти формы у Ni₃Hf

В конце концов, осталось убедиться в том, что данный сплав способен запоминать и восстанавливать форму при разных начальных деформациях. Для этого были проведены исследования при различных начальных напряжениях деформации (рис. 18 и 19).

Заключение

В результате исследований было доказано наличие мартенситного превращения и вы-

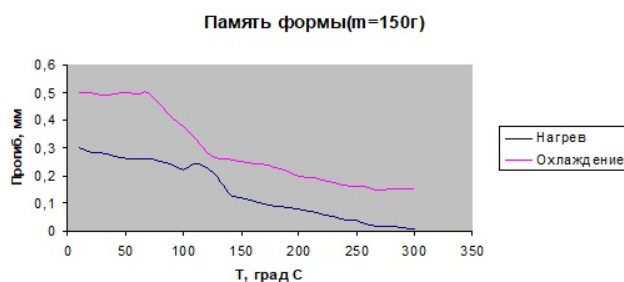


Рис. 18: Зависимость прогиба образца Ni₃Hf от температуры при нагрузке m=150 г. Степень восстанавливаемой деформации достигала 2%, а максимальное напряжение 6,3 МПа

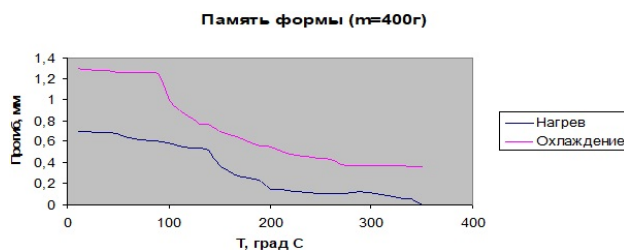


Рис. 19: Зависимость прогиба образца Ni₃Hf от температуры при нагрузке m=400 г. Степень восстанавливаемой деформации достигала 3%, а максимальное напряжение 16,7 МПа

сокотемпературного эффекта памяти формы в обоих сплавах: Ni₃Ta и Ni₃Hf. Дальнейшие шаги будут заключаться в попытке создания композитного материала на основе Ni₃Ta с добавлением Ni₃Hf в качестве легирующего компонента и исследовании свойств нового материала с перспективой повысить его износостойкость и прочность до уровня, достаточного для коммерческого применения в технике и медицине.

Литература

- [1] Лободюк В.А., Эстрин Э.И. «Мартенситные превращения», Москва, ФИЗМАТЛИТ, 2009.
- [2] Коваль Ю.Н., Фирстов Г.С. Никель-тантал — новый объект для исследования мартенситных превращений // Металлофиз. Новейшие технол. — 2007. — 29. - №6. — С. 815-821.
- [3] Оцука К., Симидзу К., Судзуки Ю. Сплавы с эффектом памяти формы: Пер. с яп. / Под ред. Х. Фунакубо. М.: Металлургия, 1990. — 224 с.
- [4] Хмелевская И.Ю., «Сплавы с эффектом памяти формы и их применение в медицине и технике», МИСиС, Москва, 2009.
- [5] Рогов В.А., Позняк Г.Г. «Современные машиностроительные материалы и заготовки», издательский центр «Академия», Москва, 2008.

- [6] «Boeing's Variable Geometry Chevron, Morphing Aerostructure for Jet Noise Reduction», The Boeing Company. Published by the American Institute of Aeronautics and Astronautics, Inc., 2006.
- [7] Ma J., Karaman I. and Noebe R.D. "High temperature shape memory alloys", Institute of Materials and Mining and ASM International, 2010.
- [8] Firstov G.S., Koval Yu.N., Van Humbeeck J., Ochin P. «Martensitic transformation and shape memory effect in Ni_3Ta : A novel high-temperature shape memory alloy», *Materials Science and Engineering A* — 2008. — P. 590-593.
- [9] Firstov G.S., Koval Yu.N., Van Humbeeck J., Ochin P., «High-temperature shape memory alloys. Some recent developments» *Materials Science and Engineering A* — 2006.
- [10] Firstov G.S., Koval Yu.N., Van Humbeeck J., Ochin P. «High temperature shape memory alloys. Problems and Prospects» *Journal of Intelligent Material Systems and Structures*, Vol. 17 — December 2006.

Разработка методов управления движением двухколесного мобильного робота*

В. И. Берзин, Н. А. Какурников

berzin@phystech.edu

Московский физико-технический институт (государственный университет)

В работе рассматриваются методы планирования и управления движением мобильного робота. Моделируется исполнительная система на примере двухколесной приводной платформы с массивом курсовых дальнометров.

Ключевые слова: *Транспортная робототехника, система автоматического управления, навигация мобильных роботов.*

The development of control system for two-wheeled mobile robot*

V. I. Berzin, N. A. Kakurnikov

Moscow Institute of Physics and Technology (State University)

The paper discusses the issues of implementing control system for transport robot. There was kinematic model of the two-wheeled mobile robot created. On the basis of this model there was control system developed.

Keywords: *Transport robotics, control systems, indoor navigation.*

1. Введение

Навигационную систему мобильного робота можно условно представить в виде трех блоков (рис. 1): блок позиционирования и планирования траектории, блок управления движением и исполнительный блок. Первый рассматривался в прошлой статье [1], где авторами была предложена система, которая позволяет роботу строить карту помещения и с высокой точностью определять свое местоположение, но при этом не затрагивалось такое важное понятие как управление движением. Именно на этапе управления осуществляется фактическая навигация: отслеживание запланированной траектории, выявление ошибок движения и их корректировка. Таким образом, построение подобных систем управления является актуальной задачей для транспортной робототехники.

На сегодняшний день уже существуют алгоритмы, заложенные в мощных программных пакетах (например ROS), которые способны решать задачи слежения и управления. Но они достаточно требовательны к аппаратному обеспечению и рассчитаны

на производительную вычислительную базу, а также продвинутую сенсорную вроде стереокамер или датчиков, работающих по принципу Kinect. Стоит заметить, что первое является следствием второго. Другими словами, необходимость в быстрых вычислениях вызвана большими массивами данных с видеокамер.

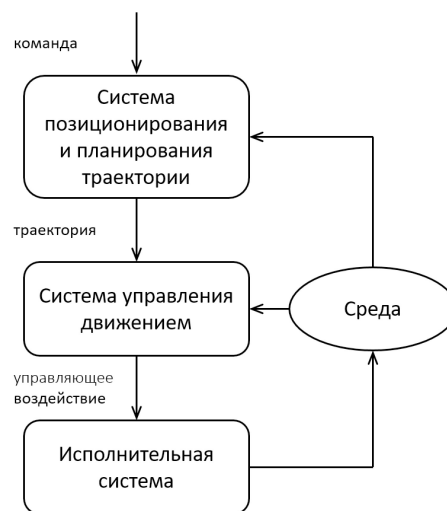


Рис. 1: Представление транспортной системы мобильного робота

Целью данной статьи является построение системы регулирования, которая могла бы полностью быть реализована на малопроизводительном встраиваемом аппаратном обеспечении (например микроконтролл-

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 16-07-00089, 18-07-20030, 18-37-10008

лерах). Ниже будет описана модель двухколесной платформы с массивом курсовых дальномеров на борту, а также предложена система автоматического управления движением по заданной траектории для мобильных роботов, подчиняющихся принципам описанной модели.

2. Постановка задачи

Пусть имеем одноосевую двухколесную мобильную платформу (рис. 2). Положение робота в декартовых координатах определим вектором:

$$\mathbf{q} = \begin{pmatrix} x \\ y \\ \varphi \end{pmatrix}.$$

Здесь $(x, y) \in XY$, φ – угол между диаметральной плоскостью робота и направлением оси абсцисс (далее курсовой угол).

Пусть ω_R, ω_L – угловые скорости вращения правого и левого колеса соответственно, r – радиус колеса, $2b$ – расстояния между центрами колес (далее база робота), v_R, v_L – линейные скорости движения центров правого и левого колеса соответственно, ω, v – угловая и линейная скорости движения центра масс робота.

Пусть также робот оснащен массивом из трех дальномеров:

$$\mathbf{d} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}.$$

Здесь d_n – текущие показания n -го дальнотмера.

Наконец, пусть робот обладает знаниями о местности (далее карта). Определим карту как отображение:

$$\mathbf{q}_i \rightarrow \mathbf{d}_i.$$

С помощью данного отображения можно спрогнозировать показания дальномеров в любой точке \mathbf{q}_i , принадлежащей карте.

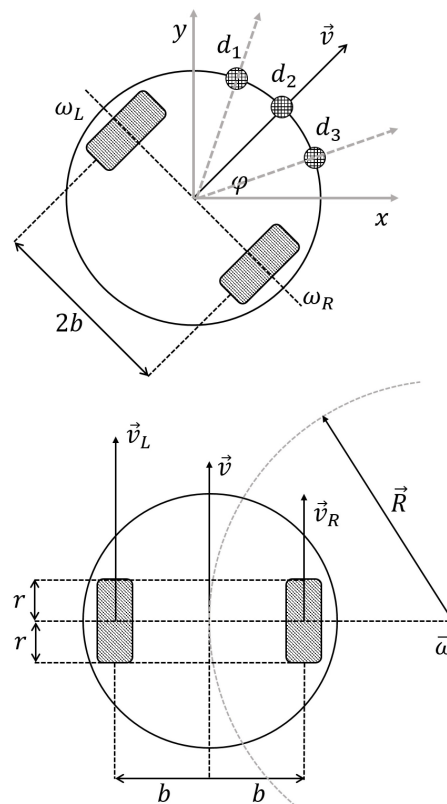


Рис. 2: Кинематическая схема робота

Рассмотрим задачу движения обозначенной платформы по заданной траектории:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}.$$

Продифференцировав по времени, получим:

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} v_x(t) \\ v_y(t) \end{pmatrix}.$$

Так как двухколесный робот не может перемещаться вбок без проскальзываний, то на него действуют кинематические ограничения. Тогда для неголономной системы имеем:

$$\begin{pmatrix} v_x(t) \\ v_y(t) \end{pmatrix} = \begin{pmatrix} v \cos \varphi \\ v \sin \varphi \end{pmatrix}.$$

Очевидно, что если робот движется по траектории с известной скоростью, то всегда известен его курсовой угол — он равен тангенсу угла наклона касательной к траектории (вектора скорости). Теперь можем за-

писать:

$$\dot{\mathbf{q}} = \begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{\varphi} \end{pmatrix} = \begin{pmatrix} v \cos \varphi \\ v \sin \varphi \\ \omega \end{pmatrix}$$

или в матричной форме:

$$\dot{\mathbf{q}} = \begin{pmatrix} \cos \varphi & 0 \\ \sin \varphi & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} v \\ \omega \end{pmatrix}. \quad (1)$$

3. Кинематическая модель робота

В формуле (1) используется линейная и угловая скорости, которые сложно измерить непосредственно. Гораздо проще отследить угловые скорости вращения колес ω_R и ω_L .

Имеем (см. рис. 2):

$$\begin{cases} v_L = \omega_L r = \omega(R - b) = \omega R - \omega b, \\ v_R = \omega_R r = \omega(R + b) = \omega R + \omega b, \\ v = \omega R. \end{cases}$$

Исключив из системы уравнений R , получим:

$$\begin{cases} \omega_L r = v - \omega b, \\ \omega_R r = v + \omega b. \end{cases}$$

Вычтем уравнения системы:

$$\omega = \frac{\omega_R - \omega_L}{2b} r.$$

Сложим уравнения:

$$v = \frac{\omega_R + \omega_L}{2} r.$$

Подставив ω и v в (1), получим:

$$\dot{\mathbf{q}} = \frac{1}{2} \begin{pmatrix} r \cos \varphi & r \cos \varphi \\ r \sin \varphi & r \sin \varphi \\ \frac{r}{b} & -\frac{r}{b} \end{pmatrix} \begin{pmatrix} \omega_R \\ \omega_L \end{pmatrix}.$$

4. Формирование управляющего воздействия

Теперь мы знаем, с какой скоростью нужно вращать колеса в каждый момент времени, чтобы двигаться по заданной траектории. Но в реальном мире существует множество факторов, которые будут сбивать робота с расчетного курса.

Поэтому одной кинематической модели здесь будет недостаточно. Начнем с того, что подав нужное управляющее воздействие на привод, мы не можем быть уверены в том, что колесо начнет вращение с нужной нам скоростью. Но если мы будем отслеживать движение через обратную связь (например с помощью энкодера вращения), то всегда сможем проверить, достигли мы нужной скорости или нет и скорректировать управляющий сигнал.

Для поддержания заданной скорости вращения удобно использовать ПИД-регулятор. На рис. 3 показано формирование управляющего воздействия. На рис. 4 можно увидеть схему ПИД-регулятора. Здесь ОУ — объект управления (приводы с энкодерами), Z^{-1} — блок, задерживающий сигнал на один период квантования, K_p — пропорциональное звено, K_i — интегральное звено, K_d — дифференциальное звено, ω — заданная угловая скорость вращения вала привода, ω_{real} — скорость, возвращаемая энкодером.

5. Удержание робота на траектории

Мы можем задавать нужную скорость вращения вала приводов. Тем не менее существует множество других факторов, которые будут вносить свой вклад в ошибку локализации. Так например, расчет движения

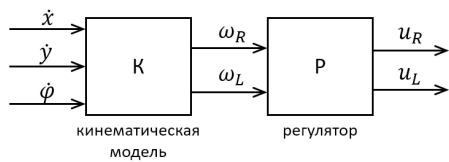


Рис. 3: Формирование управляющего воздействия

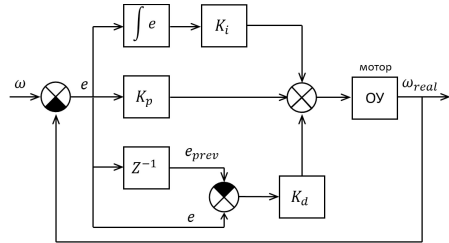


Рис. 4: Схема скоростного ПИД-регулятора

робот производит с некоторой дискретностью, на показания датчиков накладываются шумы, а колеса могут проскальзывать. Все это постепенно отклоняет робота от запланированной траектории.

Для того, чтобы предотвращать дальнейший сход с траектории, нужно вовремя детектировать отклонение и правильно на него реагировать. Другими словами, нужно добавить в систему еще один регулятор, который будет отслеживать прохождение траектории роботом и в случае необходимости корректировать маршрут.

Схему предлагаемого регулятора можно увидеть на рис. 5. Здесь после очередного формирования управляющего воздействия происходит расчет координат $\tilde{\mathbf{q}}$ по показаниям гироскопа и акселерометра. Но эти координаты не являются достоверными, так как при интегрировании величин a_x, a_y интегрируется и ошибка измерения.

Поэтому рассчитанные координаты следует проверить. Так как робот имеет карту, то он может поставить в соответствие найденному вектору $\tilde{\mathbf{q}}$ вектор $\tilde{\mathbf{d}}$, то есть предсказать, какие показания дальномеров ожидаются в точке $\tilde{\mathbf{q}}$.

С другой стороны, мы можем снять реальные показания дальномеров \mathbf{d} в текущий момент времени. Сравнивая векторы \mathbf{d} и $\tilde{\mathbf{d}}$ между собой, сделаем вывод о том, оказался ли наш прогноз верным. В случае по-

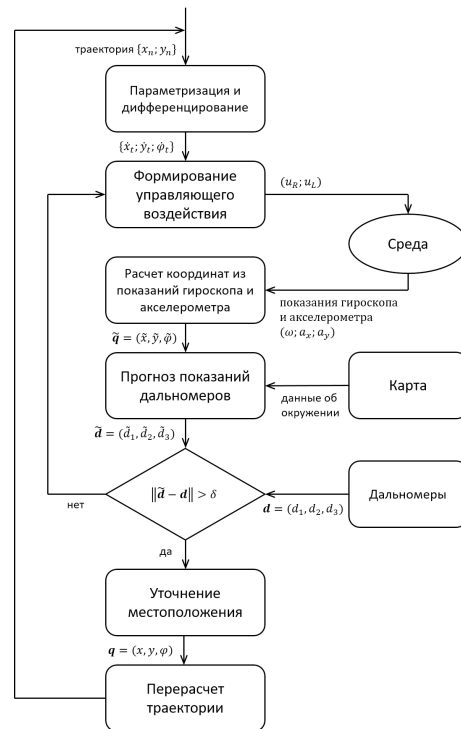


Рис. 5: Защитный триггер от схода с траектории

ложительного исхода продолжим формировать управляющее воздействие, перейдя на следующую итерацию. В противном случае сработает триггер и программа перейдет в фазу уточнения координат.

6. Уточнение координат

Уточнить координаты и скорректировать маршрут можно поручить системе позиционирования (рис. 1), но это можно сделать и другим способом значительно быстрее прямо во время движения.

Во-первых, известно, что со времени срабатывания триггера проходит небольшое время t , а скорость движения робота равна v . Тогда путь, который успеет пройти робот за это время примерно равен vt . Во-вторых, в момент срабатывания триггера $\|\tilde{\mathbf{d}} - \mathbf{d}\| > \delta$. Здесь δ имеет ту же метрику, что и пройденный путь. Тогда, вероятно, что к началу работы алгоритма уточнения координат робот не уедет дальше радиуса поиска $R_f = vt + k\delta$, где k – настраиваемый параметр.

Тем самым мы ограничили зону поиска текущих координат окружностью радиуса R_f

(рис. 6). Поиск осуществляется путем минимизации нормы $\|\tilde{\mathbf{d}} - \mathbf{d}_i\|$ для векторов $\tilde{\mathbf{q}}$ и \mathbf{q}_i в круге $\tilde{\mathbf{q}} + \mathbf{R}_f$.

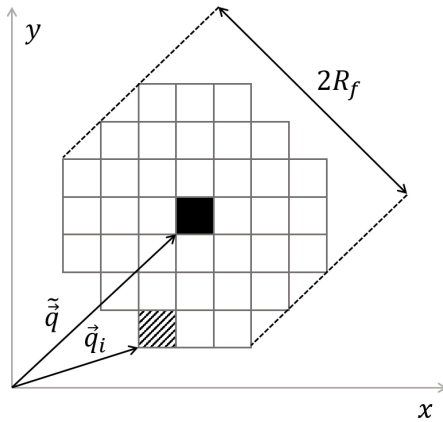


Рис. 6: Поясняющее изображение к алгоритму уточнения координат

7. Заключение

Была произведена симуляция работы некоторых методов из изложенных выше на платформе .NET(C Sharp). Алгоритмы показали высокую скорость работы и хорошую точность при уточнении местоположения. В скором времени планируется перенос программного кода на микроконтроллер и проведение испытаний мобильного робота на полигоне. Авторы надеются, что результатам эксперимента будет посвящена отдельная статья.

Литература

- [1] Берзин В.И., Судейкин М.И. Разработка методов компьютерного зрения для навигации мобильных роботов // Труды международной научной конференции СРТ1617. — 2017.
- [2] Jens Christian Andersen. Mobile Robot Navigation // PhD thesis / Technical University of Denmark. — 2006.

Выявление проблем портала государственных услуг на основе анализа текстов обращений пользователей*

А. С. Башков

bashkov@phystech.edu

Нижегородский Государственный Архитектурно-Строительный Университет

Портал государственных и муниципальных услуг является сложной, многоуровневой и высоконагруженной информационной системой. Количество его пользователей составляет более 70 млн граждан. Ежедневно службой поддержки портала обрабатываются тысячи обращений от пользователей. Мониторинг и анализ этих обращений позволяет выявлять проблемы в работоспособности и удобстве использования портала и в дальнейшем планировать работы по улучшению портала. В работе исследованы существующие методы анализа текстов и продемонстрированы результаты их работы на текстах обращений пользователей.

Ключевые слова: Электронное правительство, портал государственных услуг, анализ текстов, семантический анализ, нейронные сети

Portal of public services problems identification based on the analysis of user requests*

A. S. Bashkov

Nizhny Novgorod State University of Architecture and Civil Engineering

The portal of public services is a complex, multi-level and highly loaded information system. The number of its users is almost 70 million citizens. Every day thousands of requests from users are processed by the YSU support service. Monitoring and analysis of these calls allows you to identify problems in the operability and usability of the portal and in the future plan to improve the portal. The paper investigates existing methods of text analysis and demonstrates the results of their work on the texts of users' requests.

Keywords: e-government, government services portal, text analysis, semantic analysis, neural networks

Введение

Суть любого проекта электронного правительства, реализуемого в стране – это всегда внедрение корпоративной информационной системы национального масштаба. Ключевым различием взаимодействия бизнеса и государства с гражданами является их целевая ориентация. Для достижения эффективности бизнес ориентируется на потребности заказчика, в то время как большинство традиционных государств ориентируются на собственные нужды, в некоторой степени игнорируя потребности граждан. Ценность проектов электронного правительства заключается именно в конкретной пользе для граждан и бизнеса – в сокращении издержек, в сокращении затрат времени граждан, времени и денег бизнеса и в той добавленной стоимости, которые дают эти системы для общества.

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 17-07-00762 и 18-07-20030.

Следовательно, проект электронного правительства будет успешным тогда и только тогда, когда станет приносить реальную пользу гражданам, обществу и государству.

Для обеспечения взаимодействия граждан и бизнеса с государством был создан Единый портал государственных и муниципальных услуг. Портал представляет собой информационную систему государственного масштаба и ориентирован на пользователей всех возрастов и социальных категорий. На текущий момент их число составляет порядка 70 млн человек. В целях консультирования и помощи пользователям работает техническая поддержка. За первый квартал 2018 года в техподдержку поступило порядка 900 тысяч обращений, то есть в среднем около 10 тысяч обращений в сутки.

В целях предоставления гражданам качественного сервиса встала задача проведения анализа текстов обращений пользователей для выявления проблем портала и их

последующего устранения. Во-первых, такой анализ позволит выявлять имеющиеся массовые или вновь возникающие проблемы неработоспособности отдельных функций и сервисов. Портал обеспечивает маршрутизацию заявлений граждан к ведомствам, отслеживание статуса и получение ответа. Поэтому зачастую проблемы могут быть не на стороне портала, а на стороне ведомств, что можно узнать из реакции пользователей. Во-вторых, обратная связь от пользователей позволит установить проблемные места портала в части удобства использования (usability), исправить их или дать необходимые подсказки. В-третьих, анализ обращений позволит узнать об ожиданиях и потребностях граждан, которые впоследствии могут быть реализованы на портале.

Таким образом, анализ текстов обращений пользователей позволит повысить качество и удобство сервисов портала.

Методы анализа текстов

Автоматизированный анализ текстов развивается высокими темпами в связи с ростом количества электронных документов, статей, текстов комментариев и иных всевозможных носителей текстовой информации. Обилие такой информации делает актуальной задачу ее упорядочивания. Наиболее перспективный подход для решения этой задачи – применение технологий машинного обучения.

На текущий момент существует множество методов автоматического анализа текстов, при этом большинство из них можно разделить на три группы по подходам к анализу текста: вероятностный подход [1, 2], использование искусственных нейронных сетей [3, 4], использование деревьев решений [5].

Поскольку в большинстве случаев такие методы актуальны для решения задач аналитики большого объема данных, в том числе и big data, к методам извлечения знаний из текстов предъявляются требования

эффективности и масштабируемости. Дополнительно необходимо учитывать, что зачастую первичные текстовые данные могут содержать большое количество лишней и не значащей для анализа информации («шум»), что так же является проблемой, требующей решения.

Далее рассмотрим ключевые модели представления знаний.

Векторная модель представления знаний

Одной из ключевых проблем в задачах автоматизированного анализа текстов является проблема извлечения «сути» из текста на естественном языке и представление ее в удобной для машинной обработке форме. Векторная модель для этой задачи является одной из самых известных и простых форм представления знаний.

Впервые векторное представление знаний было представлено Джерардом Салтоном. Этот способ представления знаний был разработан для системы поиска информации SMART [6]. Формальное представление и алгоритм формирования данной модели был впервые опубликован в статье [13]. Вектора и ранее использовались в системах автоматического анализа текстов. Однако новизна этой модели заключалась в том, что в качестве компонент векторов впервые стали использоваться частоты вхождений термов в документ из коллекции. Термом – это значимое слово предметной области, обработанное с помощью стеммера Портера. Автоматическое извлечение термов из текстовой коллекции с помощью методов машинного обучения подробно рассмотрено в статье [8].

Концепция этой модели заключается в представлении каждого текста из коллекции в качестве точки в пространстве или вектора в соответствующем векторном пространстве. Тексты, которые в норме этого пространства находятся ближе друг к другу, считаются семантически схожими. Такие модели, например, используются при

формировании поискового запроса пользователя или при формировании для него персональных рекомендаций по заданному набору, в этом случае сам запрос или начальный набор задается в виде точки (вектора) этого пространства [9].

Векторная модель – это представление коллекции текстов в информационном поиске векторами из одного общего для всей коллекции векторного пространства [10]. В простейшем случае векторная модель устанавливает соответствие между каждым документом частотного спектра слов и соответствующего вектора в лексическом пространстве. В процессе поиска частотный портрет запроса рассматривается как вектор в том же пространстве и по степени близости (расстоянию или углу между векторами) определяются наиболее релевантные документы. В более продвинутых векторных моделях размерность пространства сокращается отбрасыванием наиболее распространенных или редко встречающихся слов, увеличивая тем самым процент значимости основных слов [11]. Имея такую модель для всех текстов, можно, например, находить расстояние между точками пространства и тем самым решать задачу подобию этих текстов: чем ближе расположены точки, тем более сходны эти тексты.

На основе векторного представления могут быть решены некоторые проблемы обработки текстовой информации, в частности:

- сокращение объема исходной информации для выполнения процедур анализа текста и формирования систем и баз знаний;
- синтез текста с использованием информации, извлекаемой из баз знаний.

Терм-документная матрица

Терм-документная матрица представляет собой математическую матрицу, которая описывает частоту терминов, из коллекции текстов или документов. В терм-документной матрице строки соответствуют текстам в коллекции, а столбцы соответствуют терминам. Существуют различные

схемы для определения значения каждого элемента матрицы. Одной из таких является схема *tf.idf* [12, 11]. Эти схемы актуальны в задачах обработки естественного языка, в особенности в методах латентно-семантического анализа. *tf.idf* – статистическая мера, используемая для оценки важности слова в контексте документа, являющегося частью коллекции текстов (документов) или их корпуса. Предполагается, что вес слова пропорционален количеству употребления этого слова в документе, и обратно пропорционален частоте употребления слова в других документах коллекции. В дополнение к стандартной частоте термина *tf* (*term frequency*), используемой в классических векторных алгоритмах [13], вводится специальная мера *idf* (*inverse document frequency*), указывающая на количество документов, в которых встречается каждое слово из словаря. Этот показатель позволяет сгладить результат в случае частого употребления термина в различных документах.

Семантическая сеть

Семантическая сеть – информационная модель предметной области, имеющая вид ориентированного графа, вершины которого соответствуют объектам предметной области, а дуги (ребра) задают отношения между ними. Объектами могут быть понятия, события, свойства, процессы [14, 15]. Особенность дуг между узлами заключается в том, что они имеют некоторую смысловую нагрузку, выраженную в именовании связи. Наиболее общими и часто встречающимися являются связи, обозначающие часть-целое (*part-of*), конкретный объект, класс объектов (*is-a*), подкласс-класс (*a-kind-of*, *ako*).

Для всех семантических сетей справедливо разделение по арности и количеству типов отношений [16]. По количеству типов отношений сети могут быть однородными и неоднородными. Однородные сети обладают только одним типом отношений (стрелок), например, таковой является вышеупо-

мянутая классификация биологических видов [17]. В неоднородных сетях количество типов отношений больше двух. Классические иллюстрации данной модели представления знаний представляют именно такие сети. Неоднородные сети представляют больший интерес для практических целей, но и большую сложность для исследования. Неоднородные сети можно представлять как переплетение древовидных многослойных структур. По арности:

- типичными являются сети с бинарными отношениями (связывающими ровно два понятия). Бинарные отношения очень просты и удобно изображаются на графе в виде стрелки между двух концептов. На практике зачастую требуются отношения, связывающие более двух объектов.
- N-арные. При этом возникает сложность – как изобразить подобную связь на графе, чтобы не запутаться. Концептуальные графы снимают это затруднение, представляя каждое отношение в виде отдельного узла.

По размеру [18]:

- Для решения конкретных задач, например, тех которые решают системы искусственного интеллекта.
- Семантическая сеть отраслевого масштаба должна служить базой для создания конкретных систем, не претендуя на всеобщее значение.
- Помимо концептуальных графов существуют и другие модификации семантических сетей, это является еще одной основой для классификации (по реализации).

Методы интеллектуального анализа

Наиболее простой способ анализа данных – это ручной анализ, выполняемый без применения компьютерных методов. Этот метод крайне трудоемок и затратный во всех смыслах, кроме того, его невозможно применять для работы с большими объемами данных, в особенности при наличии временных ограничений.

Первые автоматизированные способы решения задачи анализа данных заключались в задании правил и выражений, при помощи которых может проводиться классификация информации. Такое выражение может быть, к примеру, таким: «если текст содержит слова «налоги» или «штрафы», его необходимо отнести в категорию оплаты». При наличии знаний о предметной области можно составить ряд правил, которые затем автоматически применяются к поступающим текстам для их классификации. Этот подход лучше ручного, поскольку автоматизация позволит обрабатывать всю поступающую информацию. Вместе с тем, в этом случае важно на постоянной основе обеспечивать полноту и актуальность правил.

При автоматическом анализе информации набор правил и критерии принятия решений вычисляются автоматически, обучая классификатор стандартными общепринятыми словами, фразами или количественной оценкой. Однако такой подход в любом случае требует задание начального набора упорядоченной информации – ручной разметки, которая заключается в присвоении тексту или отдельной его части класса, ранга или важности. Разметка более простая задача, чем написание правил. Кроме того, разметка может формироваться в регулярном режиме. Так, например, в сервисах электронной почты при отправке писем в спам [19], мы формируем обучающую выборку для соответствующего классификатора. Таким образом, классификация текстов методами машинного обучения, основана во многом на задаче первоначальной обучающей выборки [2].

Рассмотрим несколько подходов реализации метода обучения классификатора.

Латентно-семантический анализ

Метод латентно-семантического анализа (ЛСА) позволяет выявлять значения слов с учетом контекста их использования путем обработки большого объема текстов. Модель представления текста, используемая в латентно-семантическом анализе, во

многим схожа с восприятием текста человеком. Например, с помощью этого метода можно оценить текст на соответствие заданной теме. В качестве исходной информации используется терм-документная матрица. Терм-документная матрица – это математическая матрица, описывающая частоту терминов, которые встречаются в коллекции документов. Строки соответствуют документам в коллекции, а столбцы соответствуют терминам. К матрице применяется сингулярное разложение. Сингулярное разложение – это математическая операция, раскладывающая матрицу на 3 составляющих. Сингулярное разложение можно представить в виде формулы:

$$A = USV^T,$$

где – исходная матрица, U и V^T – ортогональные матрицы, а S – диагональная матрица, значения, на диагонали которой называются сингулярными коэффициентами матрицы A . Сингулярное разложение позволяет выделить ключевые составляющие исходной матрицы. Основная идея ЛСА состоит в том, что если в качестве матрицы A использовалась терм-документная матрица, то матрица A^* , содержащая только k первых линейно независимых компонент, отражает основную структуру различных зависимостей, присутствующих в исходной матрице. Структура зависимостей определяется весовыми функциями термов. Как правило, выбор k зависит от поставленной задачи и подбирается эмпирически. Если выбранное значение слишком велико, то метод теряет свою мощьность и приближается по характеристикам к стандартным векторным методам. Слишком маленькое значение не позволяет улавливать различия между похожими терминами или документами. Если же необходимо выбирать значение k автоматически, то можно, например, установить пороговое значение сингулярных коэффициентов и отбрасывать все строки и столбцы, соответствующие сингулярным коэффициентам, не превышающим данное пороговое значение. Схожесть между любой комбинацией термов и/или доку-

ментов чаще всего вычисляют с помощью скалярного произведения их векторов, однако на практике лучший результат дает вычисление схожести с помощью коэффициента корреляции Пирсона [5].

ЛСА отображает тексты и отдельные слова в так называемое «семантическое пространство», в котором и производятся все дальнейшие сравнения. При этом делаются следующие предположения:

1. Текст – это просто набор слов. Порядок слов в текстах игнорируется. Важно только то, сколько раз то или иное слово встречается в тексте.
2. Семантическое значение текста определяется набором слов, которые, как правило, идут вместе.
3. Каждое слово имеет единственное значение. Это, безусловно, сильное упрощение, но именно оно делает проблему разрешимой.

Деревья решений

Деревья решений – один из простейших методов машинного обучения. Это совершенно прозрачный способ классификации наблюдений, и после обучения они представляются в виде последовательности предложений if-then (если-то), организованных в виде дерева [5]. Имея дерево решений, нетрудно понять, как оно принимает решения. Достаточно проследовать вниз по дереву, правильно отвечая на вопросы, – и в конечном итоге ответ будет получен. Обратная трассировка от узла, в котором произошла остановка, до корня дает обоснование выработанной классификации [21, 10].

Процесс обнаружения знаний

Для обнаружения знаний в данных недостаточно просто применить методы интеллектуального анализа, хотя, безусловно, этот этап является основным в процессе интеллектуального анализа. Весь процесс состоит из нескольких этапов:

- понимание и формулировка задачи анализа;

- подготовка данных для автоматизированного анализа (препроцессинг);
- применение методов интеллектуального анализа и построение моделей;
- проверка построенных моделей;
- интерпретация моделей человеком.

На первом этапе выполняется осмысление поставленной задачи и уточнение целей, которые должны быть достигнуты методами интеллектуального анализа. Важно правильно сформулировать цели и выбрать необходимые для их достижения методы, так как от этого зависит дальнейшая эффективность всего процесса.

Второй этап состоит в приведении данных к форме, пригодной для применения конкретных методов интеллектуального анализа. При этом вид преобразований, совершаемых над данными, во многом зависит от используемых методов, выбранных на предыдущем этапе.

Третий этап – это собственно применения методов интеллектуального анализа. Сценарии этого применения могут быть самыми и различными и могут включать сложную комбинацию разных методов, особенно если используемые методы позволяют проанализировать данные с разных точек зрения.

Следующий этап – проверка построенных моделей. Очень простой и часто используемый способ заключается в том, что все имеющиеся данные, которые необходимо анализировать, разбиваются на две группы. Как правило одна из них большего размера, другая – меньшего. На большей группе, применяя те или иные методы интеллектуального анализа, получают модели, а на меньшей – проверяют их. По разнице в точности между тестовой и обучающей группами можно судить об адекватности построенной модели.

Последний этап – интерпретация полученных моделей человеком в целях их использования для принятия решений, добавление получившихся правил и зависимостей в базы знаний и т. д. Этот этап часто подразу-

мевает использование методов, находящихся на стыке технологий интеллектуального анализа и технологии экспертных систем. От того, насколько эффективным он будет, в значительной степени зависит успех решения поставленной задачи [22].

Для подготовки данных необходимо сделать следующее.

Очистка данных – исключение противоречий и случайных «шумов» из исходных данных. При анализе любой текстовой информации имеет место быть сильная зашумленность. В связи с этим, прежде, чем переходить к анализу необходимо произвести ряд действий для освобождения текста от шумов. Стемминг – это процесс нахождения основы слова для заданного исходного слова. Основа слова необязательно совпадает с морфологическим корнем слова. Алгоритм стемматизации представляет собой давнюю проблему в области компьютерных наук. Первый документ по этому вопросу был опубликован в 1968 году. Данный процесс применяется в поисковых системах для обобщения поискового запроса пользователя [23].

Конкретные реализации стемматизации называются алгоритм стемматизации или просто стеммер. Наиболее удачный алгоритм стемминга – стеммер Портера. Стеммер Портера – алгоритм стемминга, опубликованный Мартином Портером в 1980 году. Оригинальная версия стеммера была предназначена для английского языка. Алгоритм не использует баз основ слов, а лишь, применяя последовательно ряд правил, отсекает окончания и суффиксы, основываясь на особенностях языка, в связи с чем работает быстро, но не всегда безошибочно [23].

Семантическое ядро – это подборка понятий, имеющих существенное значение для данной предметной области. Точное определение семантического ядра зависит от области применения. Так, в лингвистике, семантическим ядром называют «не упрощаемое замкнутое подмножество языка», под-

разумевая при этом скорее смысловую составляющую языка, а не грамматические конструкции. Если попытаться перейти в доступную пониманию область теории информации, оперирующую статистическими параметрами текста, то можно говорить о семантическом ядре как о подборке смысловых единиц, достаточной для классификации текста. В качестве такой единицы может выступать словоформа, лексема или другая языковая конструкция. Для использования в статистическом анализе текста можно дать определение нескольких подборок смысловых единиц, сходных с семантическим ядром, например:

1. Специфичные слова предметной области. Это такие слова, которые встречаются исключительно в текстах предметной области и позволяют установить принадлежность текста этой предметной области.
2. Высокоинформативные слова предметной области.
3. Каждое слово имеет единственное значение. Это, безусловно, сильное упрощение, но именно оно делает проблему разрешимой. Это такие слова, которые позволяют рубрицировать тексты внутри предметной области.

Семантическое ядро проще всего сформировать, анализируя большой объем текстов по предметной области. В него попадают слова, которые чаще всего встречаются в анализируемых текстах.

- Интеграция данных – объединение данных из нескольких возможных источников в одном хранилище.
- Преобразование данных. На данном этапе данные преобразуются к форме, подходящей для анализа. Часто применяется агрегация данных, дискретизация атрибутов, сжатие данных и сокращение размерности.

Обычно в системах интеллектуального анализа текстов выделяются следующие главные компоненты [24]:

1. База данных, хранилище данных или другой репозиторий информации. Это может быть одна или несколько баз данных, хранилище данных, электронные таблицы, другие виды репозитивов, над которыми могут быть выполнены очистка и интеграция. Для рассматриваемой предметной области целесообразнее всего использовать реляционных хранилища и объектно-ориентированные базы данных, поскольку оптимизированы для хранения данных подобной структуры.
2. Сервер базы данных или хранилища данных. Указанный сервер отвечает за извлечение существенных данных на основании пользовательского запроса.
3. База знаний. Это знания о предметной области, которые указывают, как проводить поиск и оценивать полезность результирующих паттернов. Данный пункт тесно взаимосвязан с предыдущим, поскольку является его частным случаем.
4. Служба получения знаний. Она является неотъемлемой частью системы интеллектуального анализа текстов и содержит набор функциональных модулей для таких задач, как характеристика, поиск ассоциаций, классификация, кластерный анализ и анализ отклонений.
5. Модуль оценки паттернов. Данный компонент вычисляет меры интереса или полезности паттернов.
6. Графический пользовательский интерфейс. Этот модуль отвечает за коммуникации между пользователем и системой интеллектуального анализа текстов, визуализацию паттернов в различных формах.

Результаты метода на данных

Для реализации метода использовались текстовые выборки из сообщений пользователей. Предметом анализа стали 858 369 обращений пользователей портала, поданные за первый квартал 2018 года. Из сообщений

масками были отсечены технические детали, персональные данные и иная информация для идентификации пользователей.

Количество символов	64 384 173
Количество слов	12 004 292
Количество уникальных слов	29 758
Количество классифицированных обращений	599 287

Было сформировано семантическое ядро поиска из выражений до 4 слов. Наибольшее число сообщений было классифицировано по теме "оплата" (штрафы, налоги, судебные задолженности, госпошлины и т.п.). Также, но с гораздо меньшей частотой, были классифицированы следующие проблемные области: подача заявления, получение квитанции по платежу, подтверждение паспортных данных, неработоспособность портала или отдельных элементов функциональности (зависание), получение паспорта, получение ВУ, запись ребенка в детский сад, получение загранпаспорта, удаление учетной записи.

По наиболее популярным проблемам, выделенным в ходе классификации, был предложен метод визуализации в виде объемной диаграммы (рис. 1).

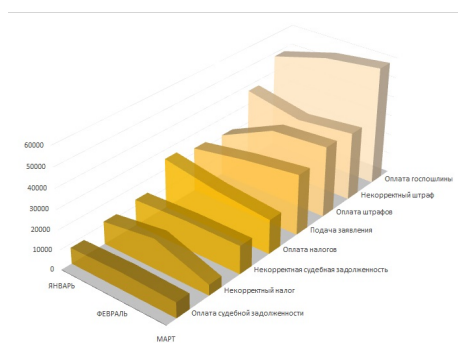


Рис. 1: Визуализация динамики поступления сообщений в техподдержку по ключевым проблемам по трём месяцам 2018 года

Заключение

Анализ текстов обращений пользователей в техническую поддержку портала позволяет в оперативном режиме отслеживать качество работоспособности портала. Внедре-

ние методов анализа в систему мониторинга портала с соответствующими механизмами визуализации позволит диагностировать массовые проблемы, ход их разрешения. Также система может выявлять изменения набора в семантическом ядре и при наборе определенного критического веса диагностировать новую возникшую проблему, что особенно актуально в условиях постоянного развития и расширения функциональности портала.

В дальнейшей работе планируется улучшать качество работы оценочных алгоритмов, а также построить автоматические средства визуализации результатов (дашборды) для руководителей, принимающих решения.

Литература

- [1] Вапник В.Н., Стерин А.М. Об упорядоченной минимизации суммарного риска в задаче распознавания образов // Автоматика и телемеханика. — 1978. — № 10. — С. 83—92.
- [2] Hastie T, Tibshirani R., Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. — Verlag : Springer, 2009. — 746 pp.
- [3] Bishop C. Neural Networks for Pattern Recognition. — Oxford : Oxford University Press, 1995. — 177 pp.
- [4] Jordan M. Mitchell T.M. Machine learning: Trends, perspectives, and prospects // Science. — 2015. — Vol. 349, no. 6245. — Pp. 255-260.
- [5] Segaran T. Programming Collective Intelligence. — LA : O'REILLY, 2008. — 368 pp.
- [6] Jaro M. A. Advances in record linkage methodology as applied to the 1985 census of Tampa Florida // Journal of the American Statistical Association. — 1989. — 84 (406). — Pp. 414-420.
- [7] Salton G., Wong A., Yang C. S. A vector space model for automatic indexing // Communications of the ACM. — 1975. — 18(11). — Pp. 613-620.
- [8] Большакова Е., Лукашевич Н., Нокель М. Извлечение однословных терминов из текстовых коллекций на основе методов машинного обучения // Информационные технологии. - 2013. - С. 31-37.
- [9] McAuley J.J., Leskovec J, Jurafsky D. Learning attitudes and attributes from multi-aspect reviews // In Proceedings of International Conference on Data Mining. — 2012. — Pp. 1020-1025.
- [10] Usama F., Smyth P., Piatetsky-Shapiro G. From Data Mining to Knowledge Discovery in Databases // Artificial intelligence Magazine. — 1996. — 17(3). — Pp. 34-54.
- [11] Teevan J. Improving information retrieval with textual analysis: Bayesian models and beyond:

- MA thesis / Teevan J.B. — Master's Thesis, Department of Electrical Engineering, Computer Science, Massachusetts Institute of Technology, 2001.
- [12] Джарратано Д, Райли Г. Экспертные системы: принципы разработки и программирование. — 4-е изд. — М. : Вильямс, 2006. — 1152 с. — ISBN 978-5-8459-1156-8.
- [13] Salton G. Improving retrieval performance by relevance feedback // Readings in information retrieval. — 1997. — Vol. 24. — Pp. 1-5.
- [14] Roussopoulos N. Conceptual Modeling: Past, Present and the Continuum of the Future // Conceptual Modeling: Foundations and Applications. — 2009. — Pp. 139-152.
- [15] Amaravadi C. S. Knowledge Management for Administrative Knowledge // Expert Systems. — 2005. — 25(2). — Pp. 53-61.
- [16] Brachman R. J. What IS-A is and isn't. An Analysis of Taxonomic Links in Semantic Networks // IEEE Computer. — 1983. — 16(10).
- [17] Helbig H. Knowledge Representation and the Semantics of Natural Language. — Berlin, Heidelberg, New York : Springer, 2006.
- [18] Quillian M. R. Semantic memory //in Semantic information processing. — 1968. — Pp. 227-270.
- [19] Павлов А., Добров Б. Обнаружение поискового спама в Вебе на основе анализа разнообразия текстов // Труды Института системного программирования РАН (электронный журнал). — 2011. — Т. 21. — С. 277–296.
- [20] Блейхут Р. Теория и практика кодов, контролирующих ошибки. - Москва : Мир, 1986. - 576 с.
- [21] Ильвовский Д. Применение семантически связанных деревьев синтаксического разбора в задаче поиска ответов на вопросы, состоящие из нескольких предложений // Научно-техническая информация. Серия 2: Информационные процессы и системы. Т. 2. — 2014. — С. 28–37.
- [22] Куприянов М., Першин А. Методика моделирования агентных поисковых систем с самовосстановлением // Известия СПбГЭТУ «ЛЭТИ». Серия «Информатика, управление и компьютерные технологии». — 2010. — С. 61–66.
- [23] Willett P. The Porter stemming algorithm: then and now // Program: Electronic Library and Information Systems. — 2006. — Vol. 40, no. 3. — Pp. 219-223.
- [24] Веретенников А. Использование дополнительных индексов для более быстрого полнотекстового поиска фраз, включающих часто встречающиеся слова // Системы управления и информационные технологии. — 2013. — 2(52). — С. 61–66.

Технология анализа слабоформализованных данных мультисенсорных систем с применением методов распределенной экспертизы*

А. А. Захарова¹, А. Г. Подвесовский²

zaa@tpu.ru | apodv@tu-bryansk.ru

¹Национальный исследовательский Томский политехнический университет

²Брянский государственный технический университет

В данной работе представлена технология адаптируемого многоуровневого анализа слабоформализованных данных мультисенсорных систем, когда источники информации характеризуются гетерогенностью, а принятие решений носит междисциплинарный характер. Рассмотрена структура процесса анализа данных на разных уровнях их представления. Для реализации отдельных этапов анализа данных предложена распределенная экспертиза, предоставляющая возможности территориально распределенного взаимодействия экспертов в разных областях, а также механизмы обеспечения согласованности и сходимости результатов экспертного оценивания.

Ключевые слова: большие гетерогенные данные, мультисенсорные системы, когнитивный анализ, принятие решений, экспертные оценки, распределенная экспертиза, согласованность экспертных суждений.

The technology of analysis of poorly formalized data of multisensor systems using the methods of distributed expertise*

A. A. Zakharova¹, A. G. Podvesovskii²

¹National Research Tomsk Polytechnic University

²Bryansk State Technical University

In this paper, the technology of an adaptable multilevel analysis of poorly formalized data from multisensor systems is presented, where information sources are characterized by heterogeneity, and decision making is interdisciplinary. The structure of the data analysis process at different levels of their representation is considered. To implement separate stages of data analysis, a distributed expertise is proposed that provides opportunities for geographically distributed interaction of experts in different fields, as well as mechanisms for ensuring consistency and convergence of the results of expert evaluation.

Keywords: large heterogeneous data, multisensory systems, cognitive analysis, decision making, expert assessments, distributed expertise, consistency of expert judgments.

Введение

В условиях цифровизации общества и всех аспектов жизнедеятельности и производства, возрастает потребность обработки больших объемов данных, которые зачастую обладают такими свойствами, как слабая структурированность, междисциплинарность и мультисенсорность [5]. Это обусловлено постоянно растущими объемами информации, гетерогенностью и асинхронностью источников ее накопления и механизмов передачи, а так же масштабностью и разноплановостью принимаемых решений по результатам обработки такой информации.

В работе предлагается технология адаптируемого анализа слабоформализованных данных мультисенсорных систем, позволяющего выполнять обработку, анализ и интерпретацию данных в условиях изменяющегося набора источников данных, а также метод оценки состояния и управления многофакторной системой на основании синхронизации и исследования независимых источников данных. В зависимости от степени формализации данных применяются визуальные, когнитивные или аналитические методы их обработки [2, 5, 6]. В качестве инструмента для реализации этапов первичной и постобработки данных предлагается распределенная экспертиза. При этом достаточное внимание уделяется вопросам обеспечения согласованности и сходимости результатов экспертного оценива-

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 18-11-00215, 18-07-20030.

ния, с целью обеспечения возможности их использования для принятия непротиворечивых и обоснованных решений. Применение технологии обсуждается на примере задачи снижения ресурсоемкости процесса проектирования месторождений нефти и газа.

Анализ данных мультисенсорных систем

Под мультисенсорными, в контексте данной работы, следует понимать системы, в которых происходит сбор, анализ и интерпретация с целью принятия решений данных, полученных из гетерогенных источников [6], которые не всегда согласуются и синхронизируются, а также носят междисциплинарный характер и при традиционном подходе подлежат экспертизе специалистами разных предметных областей. Последовательная обработка данных в такой ситуации влечет возврат и изменение решений, полученных на предыдущих этапах. Процесс анализа данных мультисенсорных систем по своей конфигурации является инвариантным относительно уровня представления данных. Различаются только методы и алгоритмы реализации его этапов. При этом при переходе от уровня к уровню увеличивается степень структурированности данных, что отражено в табл. 1.

Таблица 1: Степень структурированности данных на разных уровнях их представления

Уровень представления, данных	Степень, структурированности данных
Уровень измерений	Неструктурированные
Концептуальный уровень	Слабо структурированные
Аналитический уровень (уровень моделей/гипотез)	Структурированные
Уровень знаний	Знания (результат обработки данных)

На рис. 1 представлена схема работы с данными мультисенсорных систем, отражающая конфигурацию процесса анализа данных на каждом уровне их представления. Данная схема была построена на основе известного информационного подхода к моделированию [7] и технологии data-driven

modeling [3], но с учетом мультисенсорности, что позволило определить как специальные требования к содержанию каждого этапа процесса анализа данных, так и механизмы реализации данных этапов.

Более детально содержание и механизмы реализации этапов процесса анализа данных для разных уровней их представления приведены в табл. 2.

Итоговым результатом многоуровневого анализа мультисенсорных данных являются новые знания, которые могут быть использованы в своей деятельности исследователями и аналитиками, и для них применимы методы представления, обработки, тиражирования и использования, традиционно используемые для решения соответствующих прикладных задач.

Важным инструментом, который предлагается использовать в рамках технологии анализа данных мультисенсорных систем на ряде этапов процесса анализа, и который позволяет ускорить переход от измерений к моделям за счет снижения объема вычислений и сокращения числа итераций при вычислениях, является экспертиза получаемых результатов обработки, анализа и интерпретации данных в интерактивном режиме.

Распределенная экспертиза

Традиционно под экспертизой понимают процессы, связанные с получением и обработкой мнений экспертов — квалифицированных специалистов в соответствующей предметной области. При этом в условиях информатизации общества появляется возможность распределенного взаимодействия экспертов между собой, а также с аналитиками, лицами, принимающими решения (ЛПР) и организатором экспертизы, с использованием современных сетей связи, в первую очередь Интернет. Благодаря перечисленным обстоятельствам происходит становление нового явления, называемого сетевой экспертизой, в рамках которого создаются и активно развиваются экспертные

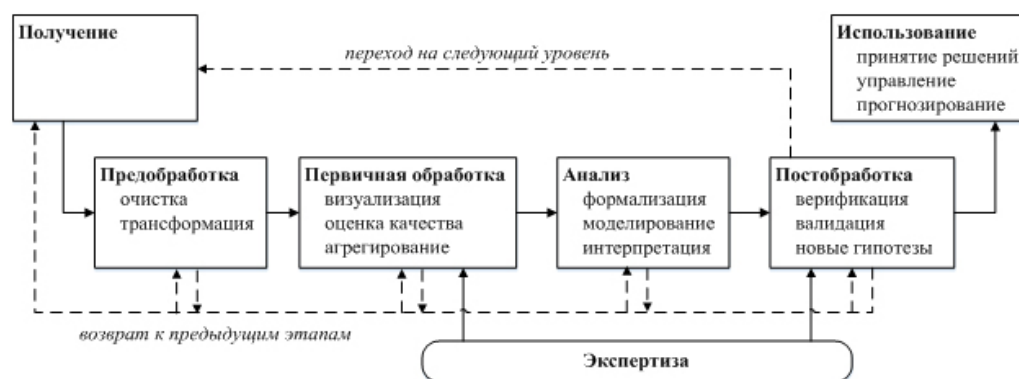


Рис. 1: Схема работы с данными мультисенсорных систем

Таблица 2: Механизмы реализации этапов процесса анализа данных для разных уровней их представления

Этап технологии	Уровень измерений	Концептуальный уровень	Аналитический уровень
Предобработка	Синхронизация; калибровка	Приведение в соответствие с требованиями модели	Устранение аномалий, группировка, фильтрация и т.п.
Первичная обработка	Визуализация многомерных данных	Построение и визуализация когнитивной карты	Подготовка данных к анализу, выбор методов и алгоритмов, обеспечение корректности их работы
Анализ	Визуальная аналитика	Структурно-целевой и сценарный анализ когнитивной карты	Статистический анализ; бизнес-аналитика; имитационное моделирование и др.
Постобработка	Снижение размерности; отбор признаков и факторов для дальнейшего анализа	Формирование и отбор гипотез для дальнейшего анализа	Представление и тиражирование знаний

сети и сетевые экспертные сообщества [1]. Об экспертных сетях говорят как о «фабриках мысли» [9] и моделях «коллективного интеллекта» [10]. Возникает новая технология анализа данных и поддержки принятия решений в распределенной среде, в рамках которой к традиционным моделям и методам анализа решений подключаются сетевые экспертные процедуры.

В то же время распределенная экспертная деятельность имеет ряд особенностей, среди которых наиболее значимыми являются изолированность экспертов друг от друга, несовпадение времени их работы и полная автоматизация получения и обработки получаемой от них информации. В связи с этими особенностями, а также с учетом ограничений по времени, усложняется задача обеспечения согласованности экспертных данных и их сходимости к некоторой общей агрегированной оценке. Таким образом, в условиях распределенной среды, эффективность традиционных методов орга-

низации экспертной деятельности решений снижается, и требуется разработка новых подходов, учитывающих специфику сетевого взаимодействия экспертов, и информационных технологий, обеспечивающих поддержку данного взаимодействия.

В контексте рассматриваемых процессов анализа данных мультисенсорных систем представляет интерес применение распределенной экспертизы в качестве одного из инструментов на этапах первичной и постобработки данных. Следует отметить, что данный подход может применяться на различных уровнях представления данных, при этом для каждого уровня реализуются собственные методы обработки данных [2, 5, 6] и формулируются цели его применения (табл. 3). Кроме того, в роли эксперта может выступать не только человек, как носитель знаний, но также и модель, либо измерительное устройство.

Рассмотрим формальную модель задачи экспертного оценивания объектов в рас-

Таблица 3: Цели применения распределенной экспертизы на различных этапах процесса анализа данных

Этап технологии	Уровень измерений	Концептуальный уровень	Аналитический уровень
Первичная обработка	Оценка качества данных на основе визуальной модели	Согласование структуры и параметров когнитивной карты	Согласование структуры и параметров применяемых моделей анализа данных
Постобработка	Обеспечение согласованности и непротиворечивости данных, полученных из разных источников	Предварительная оценка и отбор сформированных гипотез	Оценка качества решений с позиций предметной области (валидация)

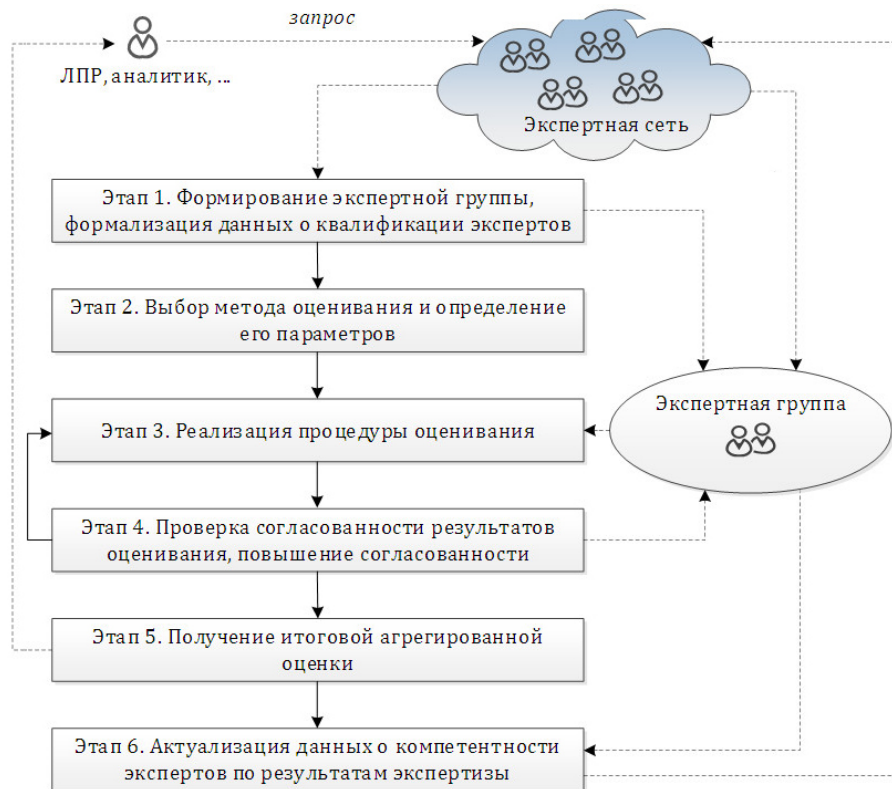


Рис. 2: Схема обобщенного алгоритма применения распределенной экспертизы

пределенной среде, построенную в рамках развития авторского подхода, предложенного в [8]. Пусть $U = \{U_1, U_2, \dots, U_p\}$ — множество предметных областей, $E = \{E_1, E_2, \dots, E_q\}$ — множество (пул) экспертов. Каждый эксперт E_s характеризуется уровнем w_{sk} квалификации (компетенций) в предметной области U_k ($s = 1, \dots, q; k = 1, \dots, p$). Данный уровень может быть выражен числовой оценкой (скалярной либо векторной), либо иметь более сложное представление. Таким образом, формируется матрица компетентности экспертов, каждая строка которой $W_s = (w_{s1}, w_{s2}, \dots, w_{sp})$ определяет профиль компетенций эксперта E_s в предметных областях, а каждый столбец $W_k =$

$= (w_{1k}, w_{2k}, \dots, w_{qk})^T$ определяет профиль охвата предметной области U_k компетенциями экспертов.

С учетом этого, задача экспертного оценивания допускает следующее формальное представление:

$$\langle T, D, X, A, F, H, C; X^* \rangle,$$

где T — тип задачи (отбор, ранжирование, оценка и др.); $D \subseteq U$ — подмножество предметных областей, с которыми связана задача; $X = x_1, x_2, \dots, x_m$ — множество объектов экспертного оценивания; A — метод (процедура или алгоритм) оценивания; F — дополнительная информация, зависящая от типа задачи T и применяемого метода оценивания A ; $H = H_1, H_2, \dots, H_n \subseteq E$

— экспертная группа (подмножество экспертов, отобранных для решения задачи); C — формализованная информация о квалификации экспертов в группе в контексте решаемой задачи; X^* — итоговое решение задачи.

Информация C играет важную роль при определении состава и численности экспертной группы H и может быть представлена в виде набора степеней относительной компетентности экспертов в предметной области D либо иметь более сложное представление. Дополнительная информация F зависит от типа задачи и, как правило, задается в форме ограничений, связанных с параметрами процедуры оценивания.

Таким образом, групповая экспертиза в распределенной среде должна обеспечивать решение таких задач, как формирование экспертной группы, выбор метода оценивания и его параметров, непосредственно оценивание, проверка согласованности результатов оценивания, определение итоговой (агрегированной) согласованной экспертной оценки. Помимо перечисленных задач возникает важная задача, связанная с оценкой компетентности экспертов в соответствующей предметной области и ее учетом в модели оценивания как на этапе формирования агрегированной оценки, так и на этапе проверки согласованности результатов оценивания. При этом, поскольку экспертную сеть можно рассматривать как группу экспертов, действующую на постоянной основе, становится возможной не только статическая, но и динамическая оценка компетентности экспертов, которая подразумевает возможность уточнения показателей компетентности каждого эксперта на основе фактических данных о его работе в группе.

Кроме того, важно учитывать, что проверка согласованности результатов оценивания не является самоцелью, а применяется для того, чтобы принять обоснованное решение о возможности использования данного множества в качестве основы при определении итоговой агрегированной групповой оцен-

ки. Тем самым, наряду с определением степени согласованности множества экспертных оценок, необходимо также оценивать ее достаточность для дальнейшей работы с данным множеством и если она окажется недостаточной, предусматривать процедуры, направленные на повышение согласованности.

Алгоритм применения распределенной экспертизы в процессе анализа данных мультисенсорных систем

Процесс применения распределенной экспертизы в качестве инструмента на этапах первичной и постобработки данных можно представить в виде обобщенного алгоритма, схема которого приведена на рис. 2. На схеме показаны этапы, соответствующие задачам, возникающим в ходе данного процесса. Процесс проходит три последовательных стадии: подготовка к экспертизе (этап 1), непосредственно проведение экспертизы (этапы 2–3) и обработка результатов экспертизы (этапы 4–6). При этом на стадии обработки результатов экспертизы ключевую роль играет этап 4, поскольку именно на нем принимается решение о возможности дальнейшего использования полученных результатов индивидуального оценивания для формирования состоятельной итоговой оценки. Следует отметить, что при этом необходимо не только проверять согласованность результатов, но и обеспечивать меры, направленные на ее повышение. Таким образом, требуется наличие механизма обратной связи, обеспечивающего возможность оперативно корректировать экспертные оценки в режиме интерактивного взаимодействия с экспертами.

С целью обеспечения программной поддержки процесса распределенной экспертизы требуется разработка комплекса математических моделей, применяемых на различных его этапах. С учетом выделенных стадий экспертизы, предлагается следующая схема моделирования (рис. 3):

- 1) построение модели формирования оптимального состава экспертной группы с учетом требований к квалификации экспертов;
- 2) разработка методов и алгоритмов реализации процедуры оценивания;
- 3) построение комплекса моделей обработки результатов экспертизы.

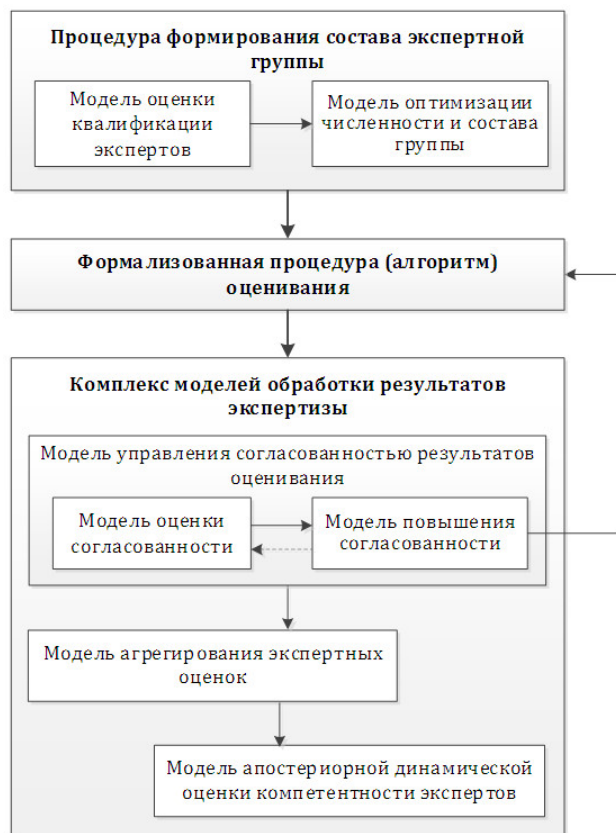


Рис. 3: Схема моделирования этапов процесса распределенной экспертизы

В рамках предложенного комплекса моделей обработки результатов экспертизы на всех уровнях представления данных наиболее важной является модель управления согласованностью результатов оценивания. Основным предназначением данной модели является обеспечение возможности использования индивидуальных экспертных оценок объектов в качестве основы для формирования итоговых групповых оценок, удовлетворяющих требованиям непротиворечивости и состоятельности. Рассмотрим ее более подробно.

Модель управления согласованностью результатов экспертного оценивания Предлагаемый подход к построению модели управления согласованностью результатов экспертного оценивания на различных этапах анализа данных мультисенсорных систем для всех уровней представления данных основан на авторском решении, предложенном в [4]. В рамках данного подхода процесс управления согласованностью предлагается представлять в виде контура с обратной связью, реализующего принцип управления по отклонению (рис. 4).

Рассмотрим общий принцип функционирования данного контура. Первым шагом после получения множества индивидуальных оценок объектов является проверка согласованности данного множества с учетом информации о компетентности экспертов. Результатом такой проверки является вычисление первичного значения коэффициента согласованности. Конкретный вид коэффициента согласованности зависит от типа используемых экспертных оценок, но общая идея, лежащая в основе вычисления данного коэффициента — оценка количества информации, содержащейся во множестве оценок, полученных от экспертов, с учетом их компетентности.

Далее необходимо установить, является ли степень согласованности множества экспертных оценок достаточной для того, чтобы его можно было использовать для вычисления агрегированной оценки. С этой целью вводится критерий оценки достаточности согласованности, основанный на некоторых специальным образом задаваемых пороговых значениях коэффициента согласованности. Сопоставление значения коэффициента согласованности с пороговыми значениями выполняется в блоке сравнения. Если согласованность множества экспертных оценок оказывается недостаточной, но при этом признается возможным ее повышение, то осуществляется переход к корректировке оценок с помощью процедуры обратной связи. В рамках данной процедуры специальным образом выби-

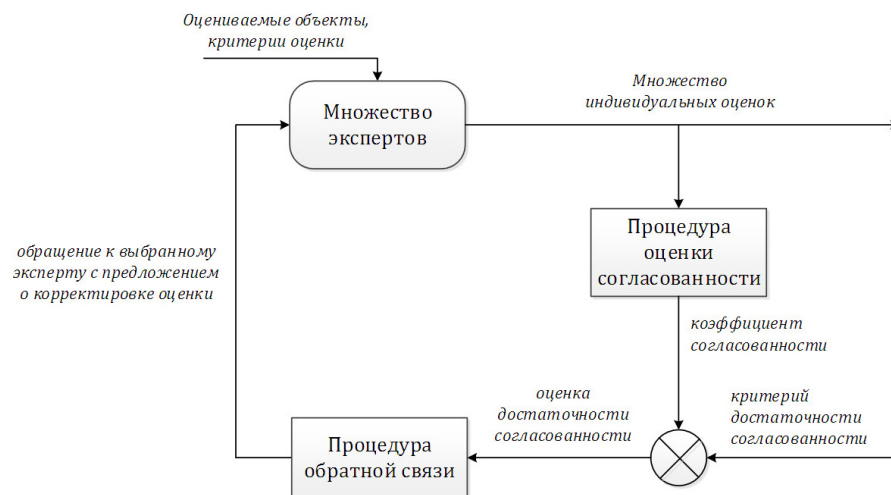


Рис. 4: Представление модели управления согласованностью результатов экспертного оценивания в виде контура с обратной связью

рается эксперт, которому предлагается изменить данную им оценку в определенном направлении (в этом и состоит управляющее воздействие). Если эксперт согласен изменить свою оценку, то по результатам изменения происходит перерасчет коэффициента согласованности. Эксперт также вправе не изменять свое суждение — в этом случае на следующем шаге будет сделана попытка выбрать другого эксперта. Процесс управления согласованностью продолжается до тех пор, пока не будет достигнут достаточный ее уровень, либо пока не будут исчерпаны возможности обращения к экспертам. В последнем случае делается вывод о невозможности получения согласованной оценки для имеющейся экспертной группы.

Предложенная модель управления согласованностью результатов экспертного оценивания может быть использована на всех уровнях представления данных, вне зависимости от степени их структурированности.

В целом предлагаемая технология многоуровневого анализа данных мультисенсорных систем с поддержкой распределенной экспертизы применима при решении широкого спектра прикладных задач в различных предметных областях. Одной из таких задач является задача проектирования месторождений нефти и газа.

Применение технологии анализа данных мультисенсорных систем при проектировании месторождений нефти и газа

Жизненный цикл проектирования месторождений нефти и газа при создании проектных документов традиционно представляет собой последовательную интерпретацию сейсмической, геофизической, геологической информации для построения статической трехмерной геологической модели. На основе такой модели, после ее адаптации на историю разработки и результаты промысловых исследований, осуществляется гидродинамическое моделирование, выполняются расчет и прогнозирование показателей экономической эффективности вариантов разработки, проектируется обустройство промысла и проводится экологическая экспертиза проекта. При этом, на каждой последующей стадии жизненного цикла осуществляется верификация проектных решений, принятых на предыдущих стадиях, с возможностью возврата на любую из предыдущих стадий, с целью корректировки параметров проектирования или расчетных статических и динамических моделей. В качестве инструмента верификации используется распределенная экспертиза, в рамках которой эксперты, каждый из которых является специали-

стом в одной или нескольких предметных областях, ассоциированных с различными стадиями жизненного цикла, осуществляют контроль и оценку согласованности и непротиворечивости параметров проектных решений. Применение технологии распределенной экспертизы, предложенной в данной работе, позволяет сократить время проектирования и снизить объем вычислений за счет повышения оперативности взаимодействия с экспертами, обеспечения возможности привлечения специалистов в необходимых областях вне зависимости от их территориального нахождения, а также наличия механизмов обеспечения согласованности и сходимости результатов экспертного оценивания.

Как показал анализ проектных решений на примере моделирования месторождений Томской области, рассмотренный подход позволяет снизить трудоемкость проектирования в 1,5–2 раза и сократить число расчетов статических и динамических моделей более чем в 2 раза.

Заключение

В работе была предложена технология адаптируемого многоуровневого анализа слабоформализованных данных мультисенсорных систем, в условиях гетерогенности источников информации, множественности охватываемых предметных областей и междисциплинарности прикладных задач, при решении которых используется данная технология. Представлена конфигурация процесса анализа данных, инвариантная относительно уровней представления данных.

В качестве подхода к реализации этапов предварительной и постобработки данных на всех уровнях их представления предложено комплексно использовать распределенную экспертизу. Представлен обобщенный алгоритм применения распределенной экспертизы, и определен комплекс математических моделей, применяемых на различных его этапах. Особенностью данного подхода является возможность территориально распределенного взаимодействия

экспертов, а также наличие механизмов обеспечения согласованности и сходимости результатов экспертного оценивания. При этом в роли экспертов могут выступать как специалисты в различных предметных областях, так и модели, либо измерительные устройства.

Применение предложенной технологии при решении прикладных задач позволяет снизить трудоемкость процесса решения и сократить объем используемых вычислительных ресурсов.

Работа выполнена при финансовой поддержке Российского научного фонда, проект № 18-11-00215.

Литература

- [1] Gubanov D., Korgin N., Novikov D., Raikov A. E-Expertise: Modern Collective Intelligence (2014) Springer International Publisher. DOI: 10.1007/978-3-319-06770-4
- [2] Isaev R.A., Podvesovskii A.G. Generalized Model of Pulse Process for Dynamic Analysis of Sylov's Fuzzy Cognitive Maps (2017) CEUR Workshop Proceedings, 1904, pp. 57-63. DOI: 10.18287/1613-0073-2017-1904-57-63
- [3] Nathan Kutz J. Data-Driven Modeling & Scientific Computation. Methods for Complex Systems & Big Data (2013) Oxford University Press, Oxford.
- [4] Podvesovskii A.G., Mikhaleva O.A. (et. al.) A Model of Control of Expert Estimates Consistency in Distributed Group Expertise (2017) A. Kravets et al. (Eds.): CIT&DS 2017, Communications in Computer and Information Science, 754, Springer International Publishing AG, pp. 361-374. DOI: 10.1007/978-3-319-65551-2_26
- [5] Zakharova A.A., Vekhter E.V., Shklyar A.V. Methods of solving problems of data analysis using analytical visual models (2017) Scientific Visualization, 9 (4), pp. 78-88. DOI: 10.26583/sv.9.4.08
- [6] Zakharova A.A., Vekhter E.V., Shklyar A.V., Pak A.J. Visual modeling in an analysis of multidimensional data (2018) Journal of Physics: Conference Series, 944 (1). DOI: 10.1088/1742-6596/944/1/012127
- [7] Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям. — СПб.: Питер, 2013. — 704 с.
- [8] Подвесовский А.Г., Михалева О.А. Особенности моделирования процессов принятия групповых решений в распределенных экспертных сетях // Вестник Брянского государственного технического университета. — 2016. — Т. 2. — № 4 (52). — С. 239-250. DOI: 10.12737/23242.

- [9] Райков А.Н. Сетевая экспертная поддержка решений / А.Н. Райков // Управление большими системами. — 2010. — № 30-1. — С. 758-773.
- [10] Славин Б.Б. Современные экспертные сети / Б.Б. Славин // Открытые системы. — 2014. — №7. — С. 30-33.

Оценка транспортного риска в проектах нефтегазовой промышленности*

Е. А. Хаританович, А. Х. Оздоева

4305@bk.ru

Москва, Российский Государственный Университет нефти и газа (НИУ) имени И. М. Губкина.

Анализ структуры инвестиции в топливно-энергетический комплекс показывает, что развитие транспортной составляющей отрасли является одним из ключевых приоритетов. Транспортные проекты в России действительно являются очень капиталоемкими. Также необходимо отметить, что реализация крупных транспортных проектов позволила сделать российский ТЭК и нефтяную отрасль более конкурентоспособными в стратегической перспективе и диверсифицированными с точки зрения структуры экспортных поставок. Это формирует масштабные возможности для развития добывающего сектора и существенно меняет потоки транспортировки как для потребителей на внешних рынках, так и в целях развития нефтеперерабатывающей промышленности России [1].

Ключевые слова: транспортный риск, нефтегазовая промышленность.

Estimation of transport risk in projects of oil and gas industry*

E. A. Kharitanovich, A. Kh. Ozdоеva

Moscow, Russian State University of Oil and Gas (Research University) Gubkin

Analysis of the structure of investments in the fuel and energy complex shows that the development of the transport sector is one of the key priorities. Transport projects in Russia are indeed very capital intensive. It should also be noted that the implementation of large transport projects made it possible to make the Russian fuel and energy complex and the oil industry more competitive in a strategic perspective and diversified in terms of the structure of export supplies. This forms large-scale opportunities for the development of the extractive sector and significantly changes the transportation flows both for consumers in foreign markets and for the development of the Russian oil refining industry [1].

Key words: transport risk, oil and gas industry.

Введение

Одним из значимых условий динамичного роста и устойчивого функционирования добывающих отраслей является открытие и освоение новых сырьевых регионов. Особенно это характерно для нефтяной промышленности, развитие которой на современном этапе сопряжено с постепенным истощением существующей сырьевой базы и поиском новых направлений, что позволит диверсифицировать направления поставок энергоресурсов и создаст высокоэффективные экспортные маршруты. А также создаваемые трубопроводные мощности смогут обеспечить прогнозируемый рост нефтедобычи в перспективных российских регионах, и позволить снять зависимость российского экспорта от транзита по территории сопредельных государств, повышая

роль России в обеспечении международной энергетической безопасности [2].

Однако расширение географии деятельности, помимо стратегических преимуществ, обуславливает и ряд проблемных аспектов. Один из наиболее очевидных и значимых — возрастающая удаленность регионов реализации проектов от поставщиков оборудования и комплектующих, необходимых для создания ключевых объектов нефтяной промышленности. Связующим звеном и важнейшим элементом реализации любого проекта является транспорт, задача которого — обеспечение оптимального взаимодействия между производителями оборудования, топлива, материалов, и регионом, реализации инвестиционных проектов. Поэтому наличие транспортной инфраструктуры — один из определяющих факторов успешности и инвестиционной привлекательности [2].

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 16-07-00098, 18-07-20030, 18-37-10008

Однако при транспортировке нередко возникает проблема некачественных перевозок, когда имеют место потери или повреждения материально-технических ресурсов, обусловленные воздействием на них знакопеременных механических и термодинамических нагрузок.

Под проектным понимаются грузы, которые впоследствии будут использоваться в рамках строительства или реконструкции объектов. Как видно из определения, большая часть проектных грузов классифицируется как крупногабаритные и тяжеловесные, что связано со значительными затратами на выезд диагностических и ремонтных бригад к месту проведения работ в случае их повреждения. К проектным грузам относятся, в частности, трубы большого диаметра, строительная техника и другие. Такое перечисление позволяет выделить еще одну особенность проектных грузов, оказывающую значительное влияние на значимость проблемы качества перевозок — эти группы являются дорогостоящими, утрата и покупка даже одной единицы обуславливает значительные экономические потери [3].

Для бесперебойной и эффективной работы предприятий нефтяной промышленности, в частности, реализующих транспортные проекты, необходимо оценивать не только общие риски, связанные с основной деятельностью, но также принимать во внимание транспортные риски, которые, как мы увидим далее, не выделяют в отдельную категорию, но имеющие огромное влияние на основные виды рисков, присущих нефтяной промышленности.

Поэтому целесообразным будет определить, какие основные виды транспортных проектов реализуются в Российской Федерации. На основании анализа реализуемых проектов крупнейшей компании РФ ПАО «Транснефть», деятельность которой непосредственно связана с транспортными операциями, выделим основные направления проектов:

1. Строительство — расширение сети магистральных трубопроводов (ВСТО, Куюмба-Тайшет).
2. Реконструкция — увеличение пропускной способности трубопроводных систем (ВСТО, Сковородино-Мохэ).
3. Транспортировка — обеспечение поставок нефти и нефтепродуктов на НПЗ и на экспорт (Сковородино-Мохэ, Проект Юг, Север).

Для рассмотрения связей между различными группами рисков, за основу была взята представленная выше классификация, разработанная авторами для анализа проектов нефтяной промышленности.

Разделим транспортные риски инвестиционного проекта на три укрупненные группы:

1. каско — источником риска являются транспортные средства;
2. карго — источником риска являются грузы, перемещаемые с помощью транспортных средств;
3. риск несвоевременной поставки грузов.

Выявленные взаимосвязи транспортных рисков с остальными видами рисков, представлены на рис. 1 [4].

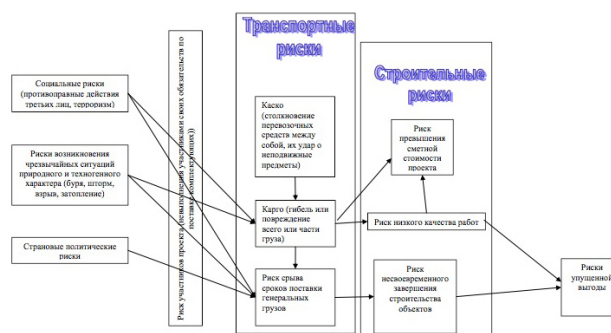


Рис. 1: Взаимосвязь транспортных рисков с другими группами рисков инвестиционных проектов

В зависимости от того, какой проект реализует предприятие, необходимо выбирать те виды транспорта, которые смогут в полной мере обеспечить его материально-техническими ресурсами.

Все проекты, непосредственно связанные с транспортировкой грузов в нефтяной промышленности можно разделить на три типа:

1. Строительство — расширение сети магистральных трубопроводов (ВСТО, Куюмба-Тайшет).
2. Реконструкция — увеличение пропускной способности трубопроводных систем (ВСТО, Сковородино-Мохэ).
3. Транспортировка — обеспечение поставок нефти и нефтепродуктов на НПЗ и на экспорт (Сковородино-Мохэ, Проект Юг, Север).

Таким образом, одной из основных задач логистики в области транспортировки является выбор транспорта. Решение данной задачи осуществляется с учетом ряда факторов объективного и субъективного характера, к которым можно отнести:

- Экономические факторы — прежде всего, это уровень затрат на транспортировку, время доставки и сохранность перевозимых грузов.
- Естественно-географические условия, обеспечивающие возможность осуществления какого-либо вида сообщения с учетом местоположения поставщиков и потребителей.
- Характер предъявляемых к транспортировке товаров и выполняемых услуг.
- Уровень развития экономических взаимоотношений партнеров, участвующих в формировании цепи поставок.
- Универсальность и специализация транспортных средств, используемых для доставки грузов.
- Степень воздействия на окружающую среду, экологическую безопасность.

В каждом конкретном случае выбора вида транспорта перечисленные выше факторы подвергаются анализу, по результатам которых и делается выбор [5].

Общая характеристика основных преимуществ и недостатков разных видов транспорта представлена в таблице 1 [5].

Логистические операции транспортировки (перемещение товаров от поставщиков до получателей, упаковка, маркировка грузовых мест, хранение товаров на промежуточных складах, сбор и разукрупнение отправок) выполняются в определенной последовательности, соответствующей технологическому транспортному процессу, в зависимости от особенностей которого выполняется выбор способа перевозки, технология доставки и конкретный маршрут. Совершенствование и появление новых технологий перевозки, модернизация объектов транспортной инфраструктуры и интеграционные процессы разных транспортных систем могут позволить снизить транспортную составляющую логистических затрат [5].

Существует несколько способов транспортировки, для их характеристики обратимся к терминологии, принятой правилами UNCTAD.

Унимодальный способ транспортировки — это перевозка грузов одним видом транспорта, при этом в данном способе транспортировки может участвовать один или несколько перевозчиков.

Интермодальный способ транспортировки — это перевозка несколькими видами транспорта с перевалкой грузов с одного вида транспорта на другой. При этом один из перевозчиков организует весь процесс доставки от одного пункта погрузки через один или несколько пунктов перевалки до конечного пункта и в зависимости от распределения ответственности между перевозчиками выдает различные виды транспортных документов.

Распространены также следующие определения способов транспортировки, различающиеся распределением ответственности между перевозчиками и видами документов.

Сегментированная перевозка — это такой способ транспортировки, который возникает, если перевозчик, организующий транспортировку, принимает ответственность лишь за часть перевозки, осуществ-

Таблица 1: Сравнительный анализ видов транспорта

№	Вид транспорта	Преимущества	Недостатки
1	Морской	<ol style="list-style-type: none"> 1. Низкие издержки на доставку. 2. Высокая производительность транспортных средств (за счет большой грузоподъемности). 3. Непрерывность работы (24 часа в сутки). 4. Незначительная зависимость от погодных условий. 5. Мобильность в зависимости от спроса. 6. Неограниченная пропускная способность морских путей. 7. Возможность перевозки любых грузов. 	<ol style="list-style-type: none"> 1. Относительно низкая скорость. 2. Сравнительно небольшая частота движения (перевозки только большого количества груза и только сразу). 3. Необходимость тщательной упаковки груза. 4. Жесткая привязка к объекту портовой инфраструктуры и зависимость от режима работы портов. 5. Сложность работы в северных районах.
2	Речной	<ol style="list-style-type: none"> 1. Высокая провозная способность речных путей сообщения. 2. Низкие издержки. 3. Возможность перевозить значительные объемы и массы грузов. 4. Возможность перевозок в районах, где нет других средств сообщения. 	<ol style="list-style-type: none"> 1. Сезонность работы на большинстве рек мира. 2. Несовпадение естественного расположения речных путей с направлением грузопотоков. 3. Необходимость строительства гидротехнических сооружений. 4. Неоднородность судоходных условий на разных реках и на отдельных участках одних и тех же рек. 5. Низкая скорость доставки.
3	Железнодорожный	<ol style="list-style-type: none"> 1. Относительно быстрая доставка на большие расстояния. 2. Независимость от климатических условий. 3. Большая грузоподъемность. 4. Сравнительно низкие тарифы. 5. Способность перевозить широкую номенклатуру грузов. 	<ol style="list-style-type: none"> 1. Зависимость от направления железнодорожного пути. 2. Необходимость перевалки грузов. 3. Переформирование составов в пути. 4. Необходимость прочной упаковки. 5. Возможность хищения перевозимых товаров.
4	Автомобильный	<ol style="list-style-type: none"> 1. Высокая маневренность. 2. Возможность адаптировать провозные возможности в зависимости от величины спроса. 3. Возможность организации срочной и регулярной доставки. 4. Сохранность грузов. 5. Экономичность при перевозках на небольшие расстояния. 	<ol style="list-style-type: none"> 1. Зависимость от дорожной сети. 2. Небольшая грузоподъемность транспортных средств. 3. Высокие тарифы. 4. Ограничения в использовании на большие расстояния.
5	Воздушный	<ol style="list-style-type: none"> 1. Высокая скорость доставки. 2. Доставка по сокращенным (спрямленным) расстояниям. 3. Высокая сохранность грузов. 4. Незначительная упаковка. 5. Высокий уровень сервиса. 6. Возможность сократить страховые запасы (из-за малого времени перевозки). 	<ol style="list-style-type: none"> 1. Самые высокие тарифы. 2. Зависимость от погодных условий. 3. Жесткие ограничения по размеру и весу. 4. Зависимость от наземных служб. 5. Удаленность аэропортов от места потребления.
6	Трубопроводный	<ol style="list-style-type: none"> 1. Независимость от погодных условий. 2. Непрерывность подачи транспортируемого груза. 3. Низкие издержки. 4. Высокий уровень автоматизации. 	<ol style="list-style-type: none"> 1. Односторонняя транспортировка. 2. Ограничение по объему груза. 3. Узкая специализация по транспортируемому грузу. 4. Высокая стоимость сооружения.

ляемой им самим. При этом он может выписать документ на интермодальную или комбинированную перевозку.

Мультимодальная перевозка — это способ транспортировки, при котором перевозчик, организующий весь процесс доставки, принимает ответственность за всю перевозку с выдачей документа о мультимодальной перевозке. Комбинированная перевозка — это перевозка грузов в одном и том же грузовом месте или транспортном средстве последовательно различными видами транспорта с использованием сквозного документа.

Смешанная перевозка — это транспортировка от пункта отправления до пункта назначения с использованием более од-

ного вида транспорта. Для организации смешанной перевозки могут привлекаться предприятия транспортной инфраструктуры (терминалы, грузовые станции и др.), либо самостоятельно, в этом случае груз последовательно передается от перевозчика одного вида транспорта перевозчику другого [5].

Таким образом, в зависимости от того, каким видом транспорта воспользуется предприятие и способа транспортировки груза, можно выделить основную нормативно-техническую и методическую документацию.

Качественный анализ рисков транспортных проектов нефтяной промышленности

Первоначальным этапом при анализе рисков инвестиционных проектов служит качественный анализ, поэтому следует уделить ему особое внимание. За последние годы теория и практика управления инвестиционными проектами обогатились различными методологиями оценки рисков инвестиционных проектов, обеспечивающие инвесторов необходимой информацией для принятия близких к оптимальным решениям. Выбор метода зависит от конкретного проекта и задач, которые ставит перед собой инвестор [6].

При этом общий алгоритм оценки проекта в условиях неопределенности и управления рисками является типичным и включает в себя следующие пункты:

1. Идентификация рисков — определение причин и источников риска, выявление этапов проекта, на которых возникает риск.
2. Количественная оценка рисков
3. Выбор методов и инструментов управления рисками.
4. Применение выбранных методов и инструментов.
5. Оценка результатов (правильности выбора методов и инструментов эффективного их использования) [6].

Основными задачами методики качественной оценки рисков проекта являются идентификация возможных видов риска проекта и их источников, определение и описание факторов и причин, влияющих на их уровень. Помимо этого, необходимо дать стоимостную оценку всех возможных последствий реализации этих рисков и определить конкретные мероприятия, позволяющие минимизировать рассматриваемый риск и/или компенсировать его последствия.

Качественный анализ проектных рисков внешне представляется очень простым,

описательным, но, по существу, его результаты служат важной исходной информацией для осуществления количественного анализа.

В методических рекомендациях ISO 31000:2009 (Risk management-Principles and guidelines) рекомендуется использовать следующие методы при оценке рисков инвестиционных проектов: SWOT анализ, анализ с помощью дерева решений, метод Делфи др. Однако, на практике, как отмечается в большинстве источников, большей степени получили применение анализ уместности затрат, экспертный метод, метод аналогий [7].

Метод аналогий

Метод аналогий заключается в анализе данных по аналогичным проектам с целью расчета вероятностей возникновения потерь.

В процессе анализа риска нового проекта используются данные о последствиях воздействия неблагоприятных факторов на другие проекты-аналоги. При выборе аналогичного проекта применяются данные исследовательских работ проектных учреждений, данные по аналогам, а также опросы менеджеров этих проектов. Далее происходит обработка собранных материалов для выявления зависимости в уже реализованных проектах для учета и оценки потенциального риска в новом инвестиционном проекте. При использовании методики необходимо сопоставлять исходные по проектам и учесть различия, так как проекты не могут быть абсолютно аналогичными.

В силу специфичности транспортных проектов нефтяной промышленности невозможно подготовить исчерпывающий набор вероятных сценариев развития проекта, которые бы реально отражали картину происходящего, так как для большинства неудач характерны свои специфические причины, которых в рассматриваемом проекте могут и не присутствовать, и аналог может не учитывать риски, реально существующие в проекте [8].

Экспертный анализ рисков

Данная методика состоит в использовании опыта экспертов в процессе анализа проекта и учета влияния различных качественных факторов. Эта методика преимущественно применяется на начальных этапах работы с проектом в случае, если исходных данных недостаточно для проведения количественного анализа эффективности проекта.

Алгоритм метода экспертного анализа риска состоит в применении следующих пунктов:

- разработка перечня критериев оценки в виде экспертных листов, содержащих вопросы;
- присвоение весовых коэффициентов для каждого критерия, неизвестных экспертной комиссии;
- составление вариантов ответов по каждому критерию, вес которых также не сообщают экспертам;
- анализ экспертами информации об оцениваемом проекте и присвоение оценок по каждому критерию;
- обработка экспертных листов статистическими пакетами и анализ результатов проведенной экспертизы.

Разновидностью метода экспертной оценки рисков, является метод Дельфи, особенностью которого является строгая процедура организации проведения оценки, при которой эксперты лишены возможности совместно обсуждать ответы на поставленные вопросы. Таким образом, максимизируется объективность в рамках данного метода.

Основным недостатком методика является трудность в подборе и привлечении независимых экспертов и субъективность оценок. Основные достоинства экспертного метода заключаются в отсутствии необходимости в наличии точной и полной исходной информации, а также необходимого программного обеспечения. Экспертный анализ рисков можно проводить до расчета параметров эффективности проекта, и он относительно прост в применении [9].

Применительно к анализу транспортного риска идентификация опасностей предполагает сбор и экспертный анализ сведений об инфраструктурных и климатических особенностях региона, где планируется осуществлять перевозки, расстоянии между поставщиком и получателем груза, количестве используемых видов транспорта и местах перевалок. На основании перечисленных данных можно делать выводы о наличии или отсутствии рисков, характерных при транспортном обеспечении определенного проекта, связанных с недостаточной развитостью инфраструктуры, сложными природно-климатическими условиями региона, мультимодальностью перевозок. Иными словами, перечисленные сведения позволяют на основании экспертных мнений сформировать начальный перечень опасностей [2].

Однако, перечисленные факторы не дают полного представления об инвестиционном проекте и не содержат достаточной информации для проведения количественного анализа. Поэтому, помимо первичных экспертных оценок, необходимо провести анализ непосредственно перевозочного процесса, в который входят:

- Инспекция транспортного средства на предмет безопасности для перевозки груза.
- Инспекция груза.
- Инспекция грузовых операций [2].

Результатом данного этапа идентификации опасностей является выявление нарушений, инцидентов, связанных с грузом, повреждений и «критических точек» на маршруте. Наиболее целесообразным и применимым, является прямая инспекция, предполагающая сопровождение груза на протяжении всего транспортного процесса, проверку перевозимых материально-технических ресурсов и транспортных средств при проведении грузовых операций, а также регистрацию всех происшествий и повреждений, имевших место в процессе транспортировки. При этом следует отметить, что распознавание транспортных рисков и их

последствий трудоемко и требует специальных знаний и достаточного опыта работы в рассматриваемой области. Оптимальным в такой ситуации является привлечение для работы независимых экспертов в области оценки рисков [2].

Таким специалистом является сюрвейер — независимый эксперт, который занимается инспектированием, исследованием, проверкой объектов страхования, а также по заявке юридических и физических лиц выдает рекомендации, связанные со страхованием рисков.

В обязанности сюрвейера входит:

- Выявление факторов риска.
- Анализ потенциальных рисков и угроз.
- Минимизация рисков путем проведения превентивных мероприятий.
- Оценка и анализ эффективности выбранного метода управления рисками.
- Оценка транспортного средства [7].

Если сюрвейер имеет дело с контролем количества и качества грузов при решении логистических задач, то он может взять на себя ответственность за определение количества и качества грузов в местах погрузки и выгрузки, а также провести оценку качества погрузочно-разгрузочных работ и проконтролировать соблюдение правил перевозки и норм хранения грузов при перевалке.

Применительно к страхованию грузов сюрвейер проводит оценку транспортных средств на предмет их пригодности для перевозки грузов, а также инспекция складских помещений и определение их соответствия требованиям и нормам хранения товаров. И в этом случае проблемы, связанные с организацией расследования причин недостачи и недоставки грузов, берет на себя сюрвейер. Помимо этого, учитывается важность контроля за количеством и качеством сырья, материалами и запчастями, которые поступают на производство, необходимость учета незавершенного производства, надзора за хранением и отпуском готовой продукции, а также за погрузочными операциями

при отгрузке готовой продукции. Все эти задачи решаются с помощью сюрвейера [7].

Так, при проведении независимой экспертизы транспортного средства, к основным задачам сюрвейера относятся:

- установление наличия и характера технических повреждений транспортного средства;
- определение причин возникновения этих повреждений;
- оценка стоимости ремонта транспортного средства.

Проведение экспертизы завершается составлением экспертного заключения, оформляемого в письменном виде (аналог «аварийного сертификата», который оформляет аварийный комиссар). В экспертном заключении должны быть указаны:

- полные координаты эксперта-техника и экспертной организации, включая гос. реестровый номер;
- дата составления и порядковый номер заключения;
- основание для проведения экспертизы;
- данные о заказчике;
- перечень и точное описание объектов, представленных для исследования и оценки в ходе экспертизы;
- нормативное, методическое и информационное обеспечение, использованное при проведении экспертизы;
- сведения о документах, рассмотренных в процессе экспертизы;
- описания проведенных исследований;
- обоснование результатов экспертизы, а также ограничения и предел применения полученных результатов;
- выводы по каждому из поставленных вопросов [7].

Таким образом, проанализировав природно-климатические условия региона, в котором планируется осуществлять перевозку, расстояние между поставщиком и получателем груза, количество используемых транспортных средств и выявив основные группы рисков, связанные с реализацией проекта, можно переходить к количественному

анализу для численного определения размеров отдельных рисков, в том числе и транспортного.

Количественный анализ рисков транспортных проектов нефтяной промышленности

Задача методов количественного анализа заключается в измерении влияния изменений факторов риска на эффективность проекта. Количественная оценка риска, то есть численное определение размеров отдельных рисков и риска проекта в целом, сложнее качественной. Сначала все риски измеряют в единицах, свойственных каждому из них, затем — в денежных единицах и, наконец, оценивают риск проекта в целом. В ряде случаев определяются вероятности наступления рисков событий и их последствий, проводится количественная оценка уровня риска, определяется допустимый в данной ситуации уровень риска [10].

В экономической литературе, посвященной проблеме риска, встречается много различных количественных методик оценки риска инвестиционных проектов. Все исследования можно условно разделить на:

1. Статистические методики.
2. Аналитические методы оценки.

К преимуществам статистических методов анализа относятся несложность математических расчетов. Недостатком данного метода является необходимость большой выборки для объективного и достоверного анализа.

Аналитические методы состоят в построении детерминированных и вероятностных аналитических моделей риска. Они включают в себя:

- метод имитационного моделирования Монте-Карло;
- метод построения «дерева решений»;
- анализ чувствительности.

Несмотря на то, что в экономических исследованиях используются различные методы

количественной оценки проектных рисков, к наиболее распространенным из них применительно к нефтяной и газовой промышленности относятся анализ чувствительности проекта, анализ сценариев развития проекта и анализ точки безубыточности.

Анализ чувствительности

Анализ чувствительности является, по сути, однофакторным анализом. В ходе анализа происходит изменение всех проверяемых на рискованность переменных: каждый раз только одна из переменных меняет свое значение на прогнозное число процентов и на этой основе пересчитывается новое значение используемого критерия. Затем все рискованные факторы ранжируются по степени влияния на изменение критерия, что позволяет выявить наиболее опасные для проекта. Базой при отборе факторов для анализа чувствительности являются результаты качественных методов [10].



Рис. 2: Чувствительность ЧДД к изменению факторов

Источник: составлено автором на основе данных ВИНК РФ

Основными целями анализа чувствительности являются:

1. Анализ влияния изменения различных факторов инвестиционного проекта на показатели эффективности проекта.
2. Выявление факторов, оказывающих наибольшее влияние на показатели эффективности проекта.
3. Определение пороговых значений данных факторов, т. е. значений, при кото-

рых инвестиционный проект становится экономически нецелесообразным [11].

Среди конкретных методов анализа чувствительности можно выделить следующие:

- *Метод критических переменных.*

Суть данного метода заключается в поиске тех значений входных параметров, которые обращают NPV в ноль (т.е. являются граничными с точки зрения принятия инвестиционного решения). По соотношению критических и наиболее вероятных значений этих параметров отбирают те, которые в наибольшей степени влияют на результаты оценки эффективности инвестиций и определяют мероприятия по предотвращению возможных негативных последствий их изменения в течение всего срока реализации проекта.

- *Метод целевых величин.*

Данный метод заключается в определении эластичности ключевых оценочных показателей эффективности к изменению различных входных параметров (например, в определении того, на сколько процентов уменьшится NPV при снижении тарифа транспортировки на 1%). Параметры с максимальной эластичностью отбираются для последующего более тщательного анализа и выбора системы мероприятий с целью минимизации рисков их значительного негативного влияния.

По сути, метод целевых величин имеет общую идеологическую основу с методом критических переменных. И в том, и в другом случаях результатом применения метода является отбор тех параметров, к изменению которых NPV (или другие показатели эффективности) наиболее чувствительны. Их отличия связаны с использованием разных алгоритмов отбора таких параметров [11].

Приведем наиболее рациональный алгоритм проведения анализа чувствительности:

1. Выбор показателей эффективности для анализа (обычно это NPV и IRR).
2. Выбор факторов, относительно которых разработчик инвестиционного проекта не имеет однозначного суждения (т.е. находится в состоянии неопределенности). Типичными факторами являются следующие факторы, свойственные транспортным проектам:
 - Тариф транспортировки;
 - Объем транспортировки (перекачки);
 - Цены на трубную продукцию и комплектующие;
 - Капитальные вложения, которые можно представить в укрупненно в виде расходов на проведение строительно-монтажных работ (СМР) и расходов на закупку материально-технических ресурсов (МТР) (оборудование и комплектующие), а также прочих расходов, осуществляемых на инвестиционной стадии, в том числе расходов на доставку МТР к месту проведения работ. Схематично полученные результаты можно представить в виде графика (рис. 2). Также не менее важным критерием эффективности проекта является дисконтированный срок окупаемости. Так как именно этот показатель показывает, какое время потребуется для того, чтобы вернуть вложенные деньги и начать получать прибыль. Схематично чувствительность срока окупаемости к изменению ключевых параметров можно представить в виде графика (рис. 3).
 - Ставка дисконтирования.
3. Установление нижних и верхних значений факторов, выбранных на втором шаге.
4. Расчет показателей эффективности при изменении выбранного фактора. Необходимо отметить, что при проведении анализа чувствительности изменяется одна переменная, а остальные фиксиру-



Рис. 3: Чувствительность срока окупаемости к изменению факторов

Источник: составлено автором на основании данных ВИНК РФ

ются на уровне базовых значений. Таким образом, выявляется влияние изменения этой переменной на показатели эффективности проекта. Для других переменных эта процедура повторяется. Расчет показателя чувствительности производится по следующей формуле [12]:

$$\Theta = \frac{\text{ЧДД} - \text{ЧДД } 100\%}{\text{ЧДД } 100\%} / \frac{X - X 100\%}{X 100\%}, \quad (1)$$

где:

Θ – эластичность;

ЧДД 100% и X 100% – базисные значения ЧДД и анализируемого показателя;

ЧДД – расчетное изменение значения ЧДД;

X – расчетное изменение значения выбранного показателя.

- Установление факторов, которые оказывают наибольшее влияние на выбранный показатель эффективности проекта [11].

Преимуществами метода анализа чувствительности критериев эффективности проекта являются — простота расчетов, объективность, теоретическая прозрачность и наглядность результатов. Помимо однофакторности, к недостаткам методики можно отнести субъективность полученных результатов из-за использования экспертного метода [13].

Анализ сценариев.

Важным преимуществом методики является тот факт, что отклонения параметров рассчитывают с учетом их взаимозависимостей (корреляции) [14]. Согласно данной методике, для рассматриваемого проекта разрабатываются несколько сценариев развития событий: базовый, оптимистический и пессимистический.

Построение пессимистического сценария связано с ухудшением значений переменных параметров до определенного разумного уровня по сравнению с базовым (реалистическим). На основании полученных значений факторов рассчитываются значения критериев эффективности проекта (ЧДД, ВНР и др.) и сравниваются. Полученные значения критериев эффективности сравниваются с их базисными значениями и формулируются необходимые рекомендации. В основе рекомендаций лежит обязательное условие: даже в оптимистическом варианте нет возможности оставить проект для дальнейшего рассмотрения, если рассчитанное значение находится за пределами эффективности проекта (например, ЧДД проекта отрицателен), и, наоборот, при пессимистическом сценарии получение, например, положительного значения ЧДД, позволяет эксперту говорить о приемлемости данного проекта [12].

Сценарный анализ позволяет инвесторам оценить доходность по каждому из сценариев и вероятность развития событий по каждому из них. Он позволяет учесть такие факторы как:

- Объем транспортировки.
- Капитальные вложения.
- Текущие затраты.
- Тариф на перекачку и др.

Анализ чувствительности и сценарный анализ являются последовательными шагами в количественном анализе рисков, при этом последний позволяет избавиться от некоторых недостатков метода анализа чувствительности. Однако, следует отметить, что

метод сценариев наиболее эффективно применим в случае, когда количество возможных значений выбранного критерия эффективности проекта конечно.

В зависимости от результатов анализа рисков, а с учетом склонности к риску инвестор, принимает решение принять, изменить или отклонить проект.

Метод Монте-Карло

Данная методика позволяет оценить риски путем создания случайных сценариев развития проекта. Результат анализа риска выражается не одним показателем значения эффективности (NPV, IRR), а в виде вероятностного распределения всех возможных значений этого показателя, т.е. инвестор будет обладать полным набором данных, характеризующих риск проекта на основе которого он сможет принять взвешенное решение о финансировании [15].

Моделирование производится с помощью специального программного обеспечения — математическая модель определения какого-либо финансового показателя подвергается ряду имитационных прогонов с помощью компьютера. Процесс имитации осуществляется таким образом, чтобы случайный выбор значений из определенных вероятностных распределений не нарушал существования известных или предполагаемых отношений корреляции среди переменных. Результаты имитации собираются и анализируются статистически для того, чтобы оценить меру риска.

Имитационное моделирование по методу Монте-Карло позволяет построить математическую модель для проекта с неопределенными значениями параметров. Зная вероятностные распределения параметров проекта, а также связь между изменениями параметров (корреляцию), можно получить распределение необходимых показателей, таких как ЧДД, ВНР и др.

Алгоритм использования метода выглядит следующим образом:

1. Определяется функция распределения каждой переменной (ключевых предпосылок), которые оказывают влияние на формирование денежных потоков. Как правило, предполагается, что функция распределения является нормальной.
2. Для каждой переменной в соответствии с ее функцией распределения генерируется случайное значение. Как правило, для генерации значений переменных используется специальное программное обеспечение.
3. Для заданных значений переменных рассчитывается ЧДД проекта.
4. Все предыдущие шаги повторяются большое количество раз, например, 10 000.
5. Полученные значения ЧДД используются для построения плотности распределения данной величины.
6. Вычисляется коэффициент вариации ЧДД проекта и затем оценивается индивидуальный риск проекта [16].

Одно из главных преимуществ метода Монте-Карло — возможность наглядно представить вероятность достижения приемлемых показателей проекта с учетом неопределенности различных факторов, заложенных в финансово-экономической модели. Основным недостатком метода является его зависимость от заложенных статистических предпосылок распределения переменных [16].

Рассмотрев основные методы качественной и количественной оценки рисков инвестиционных проектов, оценим влияние транспортного риска на эффективность инвестиционного проекта в целом.

Несмотря на существенные отличия отдельных проектов их финансовая оценка имеет единую фундаментальную теоретико-методическую основу и базируется на определенной системе показателей.

Для оценки эффективности любого инвестиционного проекта необходимо определить четыре основных критерия, на основании которых делается заключение об эф-

фективности инвестиционного проекта. К таким критериям относятся:

1. Чистый дисконтированный доход (ЧДД или NPV).
2. Внутренняя норма рентабельности (ВНР или IRR).
3. Срок окупаемости.
4. Индекс доходности.

Одним из ключевых показателей эффективности инвестиционного проекта является ЧДД. Он представляет собой разницу между всеми притоками и оттоками проекта, приведенными к текущему моменту (моменту оценки инвестиционного проекта). Он показывает величину денежных средств, которую инвестор ожидает получить от проекта после того, как денежные притоки окупят его первоначальные инвестиционные затраты и периодические денежные оттоки, связанные с осуществлением проекта.

Основными параметрами, определяющими величину ЧДД, являются капитальные вложения, эксплуатационные затраты, налоги, выручка от реализации продукции. Данное утверждение можно записать в виде функции, отражающей зависимость результирующего показателя от параметров, входящих в его расчет [12]:

$$\text{ЧДД} = F(\text{КВ}, \text{ЭЗ}, \text{В}, \text{НВ}), \quad (2)$$

где:

КВ – капитальные вложения;

ЭЗ – эксплуатационные затраты;

В – выручка от реализации продукции;

НВ – налоговые выплаты.

Организация и проведение закупок и поставок проводится на инвестиционной стадии жизненного цикла проекта [8]. Именно на этой стадии осуществляется строительство объектов, оснащение оборудованием, а также все вложения по закупке, доставке, вводу в эксплуатацию. Стоимость оборудования и комплектующих, а также затраты на их доставку, включаются в состав сметной документации по обустройству объектов, производятся на инвестиционной стадии и

включаются в состав капитальных вложений. Укрупненно капитальные вложения инвестиционного проекта можно представить следующим образом [12]:

$$\text{КВ} = f(\text{СМР}, \text{МТР}, \text{прочие}), \quad (3)$$

где:

СМР – расходы на проведение строительных и монтажных работ;

МТР – расходы на закупку материально-технических ресурсов (оборудование и комплектующие);

прочие – прочие затраты, осуществляемые на инвестиционной стадии, в том числе расходы на доставку материально-технических ресурсов к месту проведения работ.

Исходя из вышеизложенного, зависимость ЧДД от параметров его расчета можно преобразовать, представив в следующем виде [12]:

$$\text{ЧДД} = F(f(\text{СМР}, \text{МТР}, \text{прочие}), \text{ЭЗ}, \text{В}, \text{НВ}). \quad (4)$$

Таким образом, можно сделать вывод о том, что транспортные затраты и расходы на закупку материально-технических ресурсов, применяемых при строительстве объектов, являются частью денежных потоков инвестиционного проекта в составе капитальных вложений.

В общем виде, формула расчета ЧДД выглядит так [16]:

$$\text{ЧДД} = \sum \frac{CFt}{(1+i)^t}, \quad (5)$$

где:

CF – денежный поток в t – ом периоде,

i – ставка дисконтирования.

Следует отметить, что ставка дисконтирования должна быть равной рыночной стоимости капитала, например, фактической ставке процента по долгосрочным кредитам. Ставка дисконта должна отражать стоимость капитала, соответствующую возможной прибыли инвестора, которую он мог бы получить, вкладывая деньги в банк [12].

Если рассчитанный ЧДД положителен, то прибыльность инвестиций выше нормы дисконта. Проект с положительным ЧДД можно, таким образом считать приемлемым. Если ЧДД отрицателен, то прибыльность ниже стоимости капитала и от проекта нужно отказаться.

Следующим показателем, который является вторым наиболее значимым в оценке эффективности инвестиционных проектов служит внутренняя норма рентабельности (ВНР или IRR). Внутренняя норма рентабельности — это норма дисконта, при которой дисконтированная стоимость чистых поступлений от проекта равна дисконтированной стоимости инвестиций. ЧДД в таком случае равен нулю.

Обычно используется итерационная процедура расчета. В формулу расчета ЧДД подставляется некоторое начальное значение ВНР. Если ЧДД положителен, используется более высокая ставка дисконтирования. Если же ЧДД отрицателен при более высокой норме, то ВНР находится между этими двумя значениями [12]. Итак, по критерию внутренней нормы рентабельности инвестиционный проект следует принять, если ставка дисконта ниже, чем внутренняя норма рентабельности.

Следующий показатель, который мы рассмотрим — индекс доходности (ИД), который рассчитывается по формуле [12]]:

$$\text{ИД} = 1 + \frac{\text{ЧДД}}{\text{КВ}}, \quad (6)$$

где: КВ — капитальные вложения (первоначальные).

Если у проекта положительный ЧДД, то и $\text{ИД} > 1$, а если $\text{ЧДД} < 0$, $\text{ИД} < 1$. Следует отметить, что принимать инвестиционные решения только на основе ИД нельзя: у проекта с более высоким значением ИД может оказаться меньшая величина ЧДД. Кроме того, величина ИД может быть у проектов с одинаковой величиной ЧДД. Критерии ИД и ЧДД дополняют друг друга и их рекомендуется применять совместно.

И последний критерий эффективности инвестиционного проекта — срок окупаемости. Метод оценки по этому критерию очень прост в использовании, важен и понятен для практиков. Особое значение ему уделяют инвесторы, так как им необходимо знать, в течение какого периода времени окупятся его первоначальные инвестиции в проект [12].

Период окупаемости проекта определяется числом периодов планирования, в течение которых совокупные прогнозируемые потоки денежных средств покрывают первоначальные инвестиции. Обобщение основных методов оценки эффективности ИП представлено в таблице 2 [15].

Анализ преимуществ и недостатков показателей, оценивающих эффективность проекта, позволяет сделать вывод, что показатель ЧДД имеет явно выраженные преимущества по сравнению другими — внутренней нормой рентабельности, сроком окупаемости и индексом доходности. Основное его преимущество в том, что он учитывает распределение денежных потоков в течение всего срока жизни проекта и учитывает их масштаб. Однако, все рассмотренные показатели необходимо рассчитывать в проектах, ибо они дополняют общую картину эффективности проекта.

Проведение оценки влияния транспортно-го риска на экономическую эффективность проектов нефтегазовой отрасли позволяет сделать следующие выводы:

- Огромные объемы транспортируемых грузов для реализации проектов и их высокая стоимость повышают значимость своевременной и качественной доставки грузов.
- Транспортный риск находится во взаимосвязи с другими группами рисков, такими как строительные риски, риски срыва сроков реализации проекта, экологические риски и т. д.
- При недостаточном управлении транспортным риском, проекты могут стать экономически неэффективными.

Таблица 2: Обобщение основных методов оценки эффективности ИП

Показатель	Условие абсолютной приемлемости проекта	Измеритель
Чистый дисконтированный доход	$ЧДД > 0$	Ден.ед.
Индекс доходности	$ИД > 1$	Ден.ед.
Срок окупаемости	Сок 1 пр. < Сок 2 пр.	Лет

Оценка влияния транспортного риска на экономическую эффективность проекта была рассчитана на основе наиболее распространенного метода, применяемого в крупных нефтяных компаниях, анализа чувствительности. Наиболее важными факторами выступили объем материально-технических ресурсов, строительно-монтажных работ и объема транспортируемой нефти, так как при возникновении повреждений груза, в зависимости от последствий которого требуется проведение ремонта определенной сложности и стоимости или полная его замена может повлечь за собой увеличение расходов на СМР и МТР. К данным расходам относятся дополнительные расходы по заработной плате работников по монтажу/демонтажу, включая доплаты за работу в выходные дни и сверхурочные; дополнительные расходы по отчислениям на социальное страхование и т. д. Также в проектах нефтяной промышленности, связанных с транспортировкой нефти и нефтепродуктов, не редки случаи, когда происходят аварии на линейных частях магистрального трубопровода из-за брака строительно-монтажных работ, дефекта трубы и оборудования, что является прямым следствием транспортного риска. Поэтому еще одним критерием при анализе чувствительности был использован объем транспортируемой нефти. Самым серьезным фактором, который может повлечь за собой экономическую неэффективность проекта, является объем транспортируемой нефти. Однако, нефтяные компании активно внедряют новые технологии и используют методы по предотвращению этого риска, такие как установка датчиков системы обнаружения утечек нефти, создание аварийно-спасательных подразделений

на всем протяжении трассы, выбор поставщиков на тендерной основе, привлечение специалистов по анализу рисков во время транспортировки (сюрвейеров).

Литература

- [1] Соловьева Т.А., Кирьянов И.В., Сидоров А.В. Оценка и учет экономических рисков инвестиционных проектов трубопроводного транспорта нефти // Наука и технологии трубопроводного транспорта нефти и нефтепродуктов. 2011 — № 1.
- [2] Лукинский В.С. Логистика и управление цепями поставок: учебник и практикум для академического бакалавриата / В.С. Лукинский, В.В. Лукинский, Н.Г. Плетнева. — М.: Издательство Юрайт, 2016. — 359 с. — Серия: Бакалавр. Академический курс.
- [3] Трахтенгерц Э.А., Степин Ю.П., Андреев А.Ф. Компьютерные методы поддержки принятия управленческих решений в нефтегазовой промышленности. — М.: СИНТЕГ, 2005, 592 стр.
- [4] Зайцева М.А. Транспортный риск проектов газовой промышленности: диссертация — Москва, 2014. — 200с.
- [5] Щербанин Ю.А. Логистика и трейдинг в нефтегазовой отрасли. В 2 книгах. Кн. 1. Основы логистики. Закупки и поставки в нефтегазовых логистических системах: Учебное пособие. — М.: Российский государственный университет нефти и газа имени И.М. Губкина, 2015. — 280 с.
- [6] Закожурников Ю.А. Транспортировка нефти, нефтепродуктов и газа: учебное пособие для СПО / Закожурников Ю.А. — Волгоград: Ин-Фолио, 2010. — 432с.
- [7] Косов В.В. и др. Методические рекомендации по оценке эффективности инвестиционных проектов // М.: Экономика. — 2000. — Т. 2.
- [8] Марковская Е.И. Алгоритм качественной оценки рисков в проектах редевелопмента коммерческой недвижимости.
- [9] Кацал И.Н. Значение пуска нефтепровода ВСТО для экономики России. Перспективы сорта ESPO статья маркером на восточном направлении // Наука и технологии трубопроводного транспорта нефти и нефтепродуктов, — 2013 — № 1.
- [10] Зубарева В.Д., Зайцева М.А. Роль и значение транспортных рисков в системе инвестиционных проектов газовой промышленности // Научно-

- экономический журнал «Проблемы экономики и управления нефтегазовым комплексом» (Дата обращения: декабрь, 2012).
- [11] Моисеева А.В., Пахомова А.И. Анализ и оценка рисков.
- [12] Андреев А.Ф. Зубарева В.Д., Саркисов А.С., Проектные риски в нефтегазовой промышленности: Учебное пособие. — М.: «Нефть и газ», 2005. — 236 с.
- [13] Грачева М.В. Риск-менеджмент инвестиционного проекта: учебник для студентов вузов, обучающихся по экономическим специальностям. — М.: ЮНИТИ-ДАНА, 2009. — 544 с.
- [14] Теслюк Л.М., Румянцева А.В. Оценка эффективности инвестиционного проекта, 2014. — 141 с.
- [15] Сериков П.Ю. Развитие систем магистральных нефтепроводов с учетом мультипликативных эффектов межотраслевых взаимодействий: диссертация — Москва, 2017. — 370 с.
- [16] Сериков П.Ю. О рисках инвестиционных проектов строительства магистральных нефтепроводов // Наука и технологии трубопроводного транспорта нефти и нефтепродуктов. 2014 — № 1

Оценка эффективности применения технологий, направленных на снижение негативного воздействия производственных процессов нефтегазовой отрасли на окружающую среду*

А. Х. Оздоева

4305@bk.ru

Москва, Российский Государственный Университет нефти и газа (НИУ) имени И. М. Губкина

Экологическая безопасность в нефтегазовом комплексе является наиболее острой и актуальной проблемой общества и государства в целом. Под экологической безопасностью подразумевается защищенность жизни и здоровья человека, живых организмов, а также их условий обитания и произрастания от угроз природного и техногенного характера. Угрозы возникают в различных ситуациях. К примеру, на определенной территории, в результате стихийного бедствия или аварии, повлекшей за собой неминуемые жизненные и материальные потери, ущербы как здоровью человека, так и окружающей среде. К техногенным же чрезвычайным ситуациям необходимо относить различные типы аварий, таких как: пожары, взрывы, затопления, выбросы и сбросы опасных веществ и так далее.

Ключевые слова: нефтегазовая отрасль, окружающая среда.

Assessment of the effectiveness of the application of technologies aimed at reducing the negative impact of the oil and gas industry processes on the environment*

A. Kh. Ozdоеva

Moscow, Russian State University of Oil and Gas (Research University) Gubkin

Environmental safety in the oil and gas sector is the most acute and urgent problem of society and the state as a whole. Ecological safety means the protection of human life and health, living organisms, as well as their habitats and growth from natural and man-made threats. Threats arise in different situations. For example, in a certain territory, as a result of a natural disaster or an accident that caused inevitable life and material losses, damages to both human health and the environment. To the technogenic emergency situations, it is necessary to attribute various types of accidents, such as: fires, explosions, floods, emissions and discharges of hazardous substances and so on.

Keywords: oil and gas industry, environment.

Экологическая безопасность в нефтегазовом комплексе является наиболее острой и актуальной проблемой общества и государства в целом. Под экологической безопасностью подразумевается защищенность жизни и здоровья человека, живых организмов, а также их условий обитания и произрастания от угроз природного и техногенного характера. Угрозы возникают в различных ситуациях. К примеру, на определенной территории, в результате стихийного бедствия или аварии, повлекшей за собой неминуемые жизненные и материальные потери, ущербы как здоровью человека, так и окружающей среде. К техногенным же чрезвычайным ситуациям необхо-

димо относить различные типы аварий, таких как: пожары, взрывы, затопления, выбросы и сбросы опасных веществ и так далее.

Ежегодно предприятия нефтяной отрасли нарушают до 15 тыс. га земель, выбрасывают в атмосферу более 2,5 млн. тонн загрязненных веществ. Сжигают на факелах около 6 млрд. m^3 попутных газов, забирают около 750 млн. тонн пресной воды, оставляют неликвидированными сотни амбаров с буровым шламом. Отсюда и возникает актуальность и насущность проблемы безопасности окружающей среды в нефтегазовом комплексе [1, 2].

Среди проблем экологической безопасности наиболее актуальной, на мой взгляд, является охрана воздушного и водного бассейнов, так как загрязненные воздух и во-

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 16-07-00098, 18-07-20030, 18-37-10008

да обуславливают загрязнение всей биосферы Земли, включая почву и геологическую среду.

В настоящее время более эффективным методом защиты окружающей среды от загрязнения биосферы Земли является использование безотходных ресурсосберегающих и энергосберегающих технологических процессов с замкнутыми производственными циклами. Однако основными мероприятиями по защите водного и воздушного бассейнов от загрязнений выбросами вредных веществ остаются их очистка и утилизация. Несмотря на то, что подобные мероприятия остаются популярными во многих российских нефтегазовых предприятиях, разрабатываются и внедряются новые методы, техники и технологии по защите окружающей природной среды.

В федеральном законе «Об охране окружающей среды» понятие «окружающая среда» рассматривается как совокупность компонентов природной среды, которыми являются биосфера и околоземное космическое пространство, обеспечивающие благоприятные условия для существования жизни на Земле [3]. Биосфера — это единая глобальная экосистема, включающая в себя совокупность всех живых организмов на Земле — биоценоз, а также гидросферу, часть атмосферы и литосферы (табл. 1). Главной функцией биосферы является обеспечение круговорота химических элементов между гидросферой, атмосферой, литосферой и живыми организмами [4].

Таблица 1: Структура биосферы

Атмосфера — воздушная оболочка Земли, в котором только в тропосфере (до высоты 8-18 км) существуют условия, благоприятные для жизни живых организмов
Гидросфера — водная оболочка Земли, куда входят воды океанов, морей, континентальных водоемов и ледяных покровов
Литосфера — земная оболочка, где основная масса организмов сосредоточена в почвенном слое литосферы

Техносфера — среда обитания, возникающая с помощью прямого или косвенного воздействия людей и технических средств на при-

родную среду с целью наилучшего её соответствия социально-экономическим потребностям человека [5, 6].

По определению к техносфере относятся все, что создано человеком — производственная, городская, бытовая среды и т. п.

Таблица 2: Этапы и показатели развития техносферы в 20 в

Период времени и развития техносферы	Основные наименования признаков этап развития	Передовые страны
1900-1950 гг.	Электрический двигатель, ТЭС, сталь	США, Германия
1950-1980 гг.	Нефть, газ, ДВС, АЭС, авиация, космонавтика	СССР, США
1980-2000 гг.	ЭВМ	Япония, США

Создание техносферы — длительный процесс, обусловленный эволюционным развитием человека и среды его обитания. Начало активного создания техносферы приходится на начало 20 в. Рассмотрим основные признаки развития техносферы с начала и до конца 20 в. (табл. 2) [2].

В таблице 2 мы наблюдаем период научно-технической революции (НТР), в котором передовые страны создают достойную конкуренцию друг другу. Такому бурному развитию техносферы частично повлияли некоторые глобальные изменения (табл. 3) [2].

Таблица 3: Показатели глобальных изменений в 20 в

Показатели	Годы	
	1900	1990
Численность населения, млрд. чел.	1	5
Среднее потребление пресной воды одним жителем Земли за период его жизни, m^3	360	4000
Общая мощность мирового хозяйства, ТВт		1 10
Валовой мировой продукт, млрд. долл. США	60	20000
Площадь суши, нарушенная хозяйственной деятельностью, %	20	60
Средняя плотность населения, чел./кв.км	10	40
Урбанизация	13%	50%

Из приведённых в таблице 3 показателей следует, что в 20 в. на фоне демографического взрыва и урбанизации населения существенно возросло потребление продукции биоты и пресной воды, в десятки раз возросла мощность мировой экономики, в три раза увеличились освоенные человеком регионы биосферы, превратившись в города, промышленные зоны, территории АЭС, ТЭС и ГЭС, свалки и т. п. Таким образом, период НТР характеризовался несомненными достижениями в создании техносферы, реализации техники и технологий. Однако следует отметить негативные последствия для природы и человека осуществления НТР [7]:

1. Значительный рост техногенных и антропогенных опасностей, обусловленных развитием техники и технологий, очень низким уровнем адаптации человека к новым условиям жизни.
2. Увеличение влияния человека на природу.

Возникновению негативных процессов в среде обитания способствовали также просчёты государственных органов в хозяйственной деятельности, проводившейся без учёта закономерностей развития природы и потребностей человека в защите его здоровья и жизни. Примерами могут служить проект поворота вспять течения сибирских рек; строительства ГЭС, при которых в зоны затопления попадали значительные территории, занятые сельскохозяйственными угодьями.

Долгое время, вплоть до начала второй половины 20 в. человечество не замечало или игнорировало негативное воздействие от хозяйственной деятельности и техносферы на природу. О том, что человечество поздно приступило к решению проблем, связанных с негативным влиянием экономических процессов и техносферы на природную среду, свидетельствуют факты [8]:

- В США национальный закон «Об охране окружающей среды» принят в 1969 г.

- В РСФСР принят закон «об охране природы» — в 1960 г.
- В СССР закон «Об охране атмосферного воздуха» — в 1980 г.; основы водного законодательства — в 1979 г.; основы лесного законодательства — в 1977 г.
- В СССР впервые разработан и введен с 1981 г. в действие государственный стандарт «Управление охраной окружающей среды. Основные положения. ГОСТ 24525.4 — 80».
- Закон РФ «Об охране окружающей природной среды» от 1991 г.
- Федеральный закон об экологической безопасности РФ «Об охране окружающей среды» принят в декабре 2001 г. и подписан 10 января 2002 г.
- ЮНЕСКО провело конференцию по использованию и сохранению биосферы только в 1968 г. [2].

Таким образом, интенсивное развитие техносферы привело к усилению её негативного влияния на природные процессы. Факты подтверждают. Развитые и развивающиеся страны стали впервые вводить и принимать законы «Об охране окружающей среды», вводить новые понятия и термины, связанные с экологической безопасностью [2].

Поэтому весьма важно выработать понятия, адекватные сущности экологической безопасности. Экологическая безопасность — приемлемый на данном этапе социально-экономического развития степень защищённости жизненно важных интересов личности, общества, государства, мирового сообщества от последствий и угроз, которые обусловлены негативными изменениями окружающей среды, возникающими в результате антропогенного и природного воздействия на неё.

Экологические угрозы — прогнозируемые последствия или потенциальные сценарии развития событий катастрофического характера, которые обусловлены изменениями состояния окружающей среды и способны нанести вред жизненно важным интересам личности, общества, государства, ми-

рового сообщества. Обеспечение экологической безопасности — процесс формирования и реализации системного управляющего воздействия на социоэкосистему, направленный на предотвращение экологических угроз и защиту от экологических последствий вплоть до достижения приемлемого уровня защищенности. По существу, обеспечение экологической безопасности сводится к оптимизации функций полезности и безопасности [9, 10].

Защищённость объектов экологической безопасности определяется тем обстоятельством, что актуальные параметры состояния окружающей среды отклоняются от нормы, не выше предельно допустимых значений. Предельно допустимые значения актуальных параметров устанавливаются на качество компонентов окружающей среды применительно к здоровью человека, затем применительно к устойчивости естественно-природных экосистем, а также нормативы по отношению к источникам экологической опасности, которые указывают пределы действия факторов экологического риска. Степень отклонения или несоответствия актуальных параметров состояния социоэкосистемы по отношению к соответствующим нормативам определяет остроту экологических проблем и может служить критерием их приоритетности [11].

Нормативы предельно допустимых вредных воздействий на окружающую среду и здоровье человека, а также существующая система экологического контроля за их соблюдением и ответственности за их невыполнение обеспечивают нам гарантию выполнения прав граждан на здоровую и благоприятную окружающую среду, иначе эти права остались бы просто декларацией [10]. Поэтому принятые в 1990-е гг. в России принципиально новые природоохранные законы, осознание необходимости экологизации экономики и разработки действенного управления экоразвитием, начало массового экологического воспитания, образования, просвещения россиян дают надеж-

ду на оптимистическое экологическое развитие страны.

Воздействие нефтегазового комплекса на все компоненты природной среды

В деятельности нефтегазового комплекса (НГК) объектами негативного воздействия являются все компоненты ОС: атмосферный воздух, поверхностные и подземные воды, почвы, недра, растительный покров (табл. 4).

Сжигание попутного газа в факелах — это прямое загрязнение атмосферы сернистыми соединениями. В радиусе 250 м от факелов полностью уничтожается всякая растительность, на расстоянии до 3 км деревья сохнут и сбрасывают листья. Содержание в углеводородах серо- и азотосодержащие примеси сгорают с образованием оксидов серы и азота (табл. 5) [4]. Следствием выбросов в атмосферу диоксидов серы и азота являются кислотные дожди, которые оказывают разрушающее воздействие на конструкционные материалы и действуют на дыхательную систему человека.

Процессы горения оказывают влияние на климатическую систему Земли двойное воздействие: во-первых, уменьшают содержание кислорода в атмосфере, истощая озоновую защиту Земли и уменьшая атмосферное давление, во-вторых, выбрасывают огромное количество водяных паров и углекислого газа. Следствием является парниковый эффект.

При бурении скважин источниками загрязнений атмосферы являются залповые выбросы при нефте- и газопроявлениях, сжигание углеводородов на факельных установках при очистке призабойной зоны пласта, термическое обезвреживание буровых шламов [4, 12].

В период бурения скважин основными источниками выбросов в атмосферу являются дизельные установки. При аварийных разливах нефти происходит загрязнение атмосферы за счет испарения низкомолекулярных углеводородов.

Таблица 4: Воздействие нефтегазового комплекса на все компоненты окружающей среды

Виды сред	Нефтегазовый комплекс		
	Воздействие	Последствие	Причина
Атмосфера	Неативное: загрязнение атмосферного воздуха сернистыми соединениями; отрицательное влияние на флору и фауну; воздействие на климатическую систему Земли; косвенное воздействие на почвенный покров и растительность и т. д.	Неативное: развитие рака лёгких, сердечных заболеваний, инсульта и т. д.; «кислотные» дожди; парниковый эффект; уменьшение содержания кислорода в атмосфере, истощая озоновую защиту и уменьшая атмосферное давление; испарение низкомолекулярных углеводородов и т. д.	Аварии; неразвитость законодательной основы природоохранной деятельности; использование ограниченных или устаревших природосберегающих технологий; низкая экологическая культура; низкий уровень мониторинга и контроля на местах; сложный и высокочатратный процесс рекультивации; незаконные сбросы нефтешламов в сточные воды; использование буровых растворов с большим объемом воды; сжигание попутного нефтяного газа (ПНГ) на факельных установках; термическое обезвреживание буровых шламов; экономия с целью обогащения; залповые выбросы при нефте- и газопроявлениях и т. д.
Гидросфера	Неативное: загрязнение сточных вод; пагубное влияние на флору и фауну; загрязнение атмосферы путём испарения углеводородов и т. д.	Неативное: загрязнение питьевой воды; формирование анаэробных условий; угнетение и подавление нормальной органической жизни; изменение состава биоценозов; образование устойчивой водоняфтяной эмульсии; изоляция воды от кислорода воздуха; помутнение воды, вследствие, процесс фотосинтеза замедляется и т. д.	
Литосфера	Неативное: отчуждение территории; осушение или подтопление территории; вырубка леса и изменение характера землепользования; шумовые, вибрационные, электромагнитные воздействия; химическое загрязнение почв, нефтепродуктами, буровыми и тампонажными растворами; просачивание загрязнителей через почву; слив буровых сточных вод (БСВ) в покровы почв; сбросы и разливы отходов и нефтешламов и т. д.	Неативное: непригодное для использования земля в хозяйственных целях; затратный по деньгам и времени процесс рекультивации; загрязнение сточных вод; засоление грунтов; отсутствие плодородия и урожайности почв; повышение пожароопасности; испарение нефтепродуктов в атмосферу; вред флоре и фауне и т. д.	

Утилизация попутного нефтяного газа является определённой проблемой, поскольку требует сооружения на нефтепромыслах достаточно затратной дополнительной инфраструктуры — системы сбора и транспортирования попутного нефтяного газа. Из-за этого нередко можно наблюдать ситуацию, когда попутный газ просто сжигается в факелах непосредственно в районе добычи нефти. Это, безусловно, плохо не только из-за загрязнения атмосферного воздуха продуктами попутного нефтяного газа, но и потому, что данные газы являются ценным

сырьём для химической промышленности и высококалорийным топливом [13]. Поведение попавших в геологическую среду нефтепродуктов зависит от их химического состава и может сильно различаться. Некоторые фракции нефти преимущественно накапливаются в верхней части грунта и не мигрируют вглубь, другие, наоборот, способны распространяться с потоком грунтовых вод.

Источники загрязнения геологической среды нефтепродуктами можно разделить на очаговые, линейные и площадные. Наи-

Таблица 5: Серосодержащиеся примеси в углеводородах

Понятие	Определение
Оксид углерода	Токсичный газ без цвета, запаха и вкуса. Максимально допустимая концентрация (ПДК) в населенных пунктах — 3 мг/м ³
Диоксид углерода	Бесцветный тяжелый газ, вызывающий удушье. ПДК в воздухе равен 1%
Диоксид серы	Бесцветный газ с резким запахом, губительно влияющий на здоровье человека. ПДК равен 20 мг/м ³
Сажа	Вредное вещество. Разовая ПДК газовой сажи составляет 0,15 мг/м ³

большая по численности группа — это очаговые источники, к которым относятся резервуарные парки хранения нефтепродуктов, нефтеперегонные заводы, погрузочные терминалы, автозаправочные станции и другие [13].

Основные формы загрязнения грунтовых вод нефтепродуктами:

- жидкие нефтепродукты, находящиеся в верхней части слоя насыщения;
- водно-эмульсионные формы;
- водорастворимые формы углеводородов в грунтовых водах.

Наиболее опасно для населения загрязнение нефтепродуктами источников питьевого водоснабжения — скважин, колодцев, родников. С учетом того, что на территории России находятся тысячи объектов добычи, транспортировки, хранения нефтепродуктов, являющихся потенциальными загрязнителями грунтовых вод, невозможно осуществить тщательный мониторинг на всех из них. Согласно государственным докладам «О состоянии и об охране окружающей среды Российской Федерации» ежегодно в нашей стране происходят десятки крупных аварий и около 20 тысяч случаев, сопровождающихся значительными разливами нефти, загрязнением почв, подземных вод и поверхностных водных объектов, большими материальными потерями [6, 14].

При добыче и переработке нефть дважды смешивается с водой:

- при выходе с большой скоростью из скважины вместе с сопутствующей ей пластовой водой;
- в процессе обессоливания, т. е. промывки пресной водой для удаления хлористых солей.

В нефти и нефтепродуктах вода может содержаться либо в виде простой взвеси, тогда она легко отстаивается при хранении, либо в виде стойкой эмульсии, в этом случае приходится прибегать к особым приемам обезвоживания нефти.

Образование устойчивых нефтяных эмульсий приводит к большой нагрузке на поверхностные водные объекты, поскольку после отделения от нефти в отстойниках и резервуарах так называемые подтоварные воды, содержащие нефтепродукты в виде эмульсии, сбрасываются на очистные сооружения. В последующем часть нефтепродуктов улавливается и накапливается в «амбарах» и нефтяных прудах, где из неё испаряются лёгкие фракции. В итоге образуется так называемые «амбарные нефти» — высокообводнённые и смолистые соединения с большим содержанием механических примесей.

Эффективность очистки подтоварных сточных вод на существующих очистных сооружениях зачастую недостаточна, в результате реки и озёра, а также подземные водные горизонты в районах нефтепромыслов оказываются загрязнёнными нефтепродуктами.

Химическое загрязнение почв нефтепродуктами, буровыми и тампонажными растворами происходит при плохой обваловке и слабой гидроизоляции амбаров. Глубина загрязнения почвы бывает сильной — более 25 см, слабой — до 10 см. Для реанимации почв со слабой степенью загрязнённости нужно частичное снятие загрязнённого слоя, проведение вспашки в течение 2-3 лет и внесение минеральных и органических удобрений. Сильное загрязнение делает почвы непригодными для ведения сельского хозяйства. Период восстановле-

ния почвенно-растительного покрова после загрязнения нефтью в количестве 12 л/м² в зависимости от климатических особенностей может растянуться на 25 лет [15].

Происходят также аварийные разливы нефти, ведущие к серьезному нарушению биоценозов и почв на прилегающих территориях.

При добыче и подготовке к транспортировке нефти неизбежно образуются нефтесодержащие отходы, которые складываются в специальных земляных шламонакопителях — «амбарах», что приводит к загрязнению окружающей среды и потерям нефти.

Перечень отходов нефтедобычи содержит следующие наименования:

- нефтешламы;
- загрязнённые грунты, буровой шлам.

Нефтешламы — это целая группа отходов с весьма разнообразным составом, в общем являющихся многокомпонентными физико-химическими щелочесодержащими и солей металлов). По происхождению различают:

- нефтешламы, образующиеся в процессе добычи нефти;
- грунтовые нефтешламы, образующиеся в результате проливов нефтепродуктов на почву;
- природные нефтешламы, скапливающиеся на дне водных объектов при оседании в воде нефтепродуктов;
- нефтешламы резервуарного типа [12].

Таким образом, оценивая современное состояние мира опасностей, следует, к сожалению, констатировать, что в итоге эволюции человечество к концу НТР породило печальный парадокс — в течение многих лет люди совершенствовали технику, чтобы обезопасить себя от естественных опасностей, а в результате пришли к наивысшим техногенным опасностям, связанным с производством и использованием техники и технологий [16].

Технологии и методы обеспечения экологической безопасности в нефтегазовой отрасли

Для защиты окружающей среды применяют различные методы и технологии, которые снижают негативное воздействие, создаваемое нефтегазовыми процессами. Рассмотрим их подробнее на каждом виде из сред.

В качестве защиты атмосферы, применяют пылеулавливающие аппараты. Основное их назначение — это очистка газа от загрязняющих веществ. Выделяют четыре основных способа очистки газа. Каждый способ предполагает использование установок определенного типа.

1. Сухая механическая газоочистка — осаждение под действием силы тяжести и инерционных сил. Применяемые аппараты: пылесадительные камеры; инерционные пылеуловители; циклоны; электрофильтры.
2. Мокрая газоочистка — промывка загрязнённого газа водой или иными веществами. Применяемые аппараты: скрубберы; барботеры; дезинтеграторы; мокрые электрофильтры; скрубберы Вентури.
3. Фильтрация газа через пористые перегородки. Применяемые аппараты: рукавные фильтры; керамические фильтры; масляные фильтры.
4. Электрическая очистка газа — осаждение под действием электростатических сил. Применяемые аппараты: электрофильтры.

К основным характеристикам газоочистных устройств относятся: размер улавливаемой пыли; производительность устройства; гидравлическое сопротивление; расход электрической энергии; стоимость очистки [15, 17].

Размер улавливаемой пыли помогает определить эффективность очистки газа. Производительность газоочистки характеризуется количеством воздуха, которое очища-

ется за 1 час. Гидравлическое сопротивление зависит от типа аппарата и определяет необходимое давление вентилятора, а значит и расход энергии.

Стоимость очистки в различных аппаратах значительно отличается. Если стоимость очистки в циклоне принять за единицу, то стоимость очистки такого же количества воздуха в батарейном циклоне составит 1,2; в электрофильтре 2,2; в тканевых фильтрах — 2,8; в системе батарейный циклон-электрофильтр — 3,3. Выделим основные установки, которые наиболее эффективно справляются со своей работой (табл. 8).

В таблице 6 отобраны наиболее эффективные установки, которые могут с высокой степенью очистить газ от частиц пыли.

Помимо вышеперечисленных установок применяют также абсорбционные, каталитические и термические способы очистки газов.

Абсорбционная очистка основана на способности жидкостей растворять газы и химически взаимодействовать с ними. В процессе абсорбции происходит переход вещества из газовой фазы в жидкую, при десорбции, наоборот. Количество газов, которое может раствориться в жидкости, зависит от их свойств, температуры и давления газа над жидкостью. Для процесса абсорбции характерны более низкие температуры, так как при увеличении её растворимость газов в жидкости уменьшается [18].

Для очистки газов от газообразных загрязнителей применяют, например, следующие абсорбенты:

- для удаления диоксида серы — воду; суспензию известняка в воде; водный раствор соды;
- для удаления сероводорода — водный раствор фосфата калия;
- для удаления диоксида азота — воду; раствор серной кислоты или соды

Процесс адсорбции является универсальным методом, позволяющим полностью извлечь примесь из газовой или жидкой

сред. Основной задачей адсорберов является обеспечение интенсивного контакта очищаемого газа с адсорбентом. Явление адсорбции обусловлено наличием сил притяжения между молекулами адсорбента и поглощаемого вещества на границе раздела соприкасающихся фаз. Молекулы адсорбированного вещества, переходя на поверхность адсорбента, уменьшают его поверхностную энергию, в результате чего происходит выделение теплоты.

В качестве адсорбентов наибольшее распространение нашли такие материалы, как активные угли, силикагели, алюмогели и цеолиты. Хорошие адсорбенты выдерживают несколько тысяч циклов адсорбции-десорбции, что является явным преимуществом процесса адсорбции.

Следующий метод очистки газов от загрязняющих примесей — каталитический метод очистки.

Существует энергетический барьер, который нужно преодолеть частицам, чтобы они вступили в реакцию. Для снижения этого барьера используют катализаторы, которые увеличивают скорость реакций. Обычно катализатор представляет собой смесь нескольких веществ: каталитически активного вещества, активатора и носителя. В качестве каталитически активных веществ используют платину, палладий, рутений, родий и сплавы, содержащие никель, хром, ванадий, мед. В качестве носителя используют инертные пористые вещества, которые обладают развитой поверхностью: силикагели, цеолиты.

Важнейшей характеристикой катализаторов является температура зажигания — минимальная температура, при которой катализатор начинает проявлять свои свойства. В ходе катализа катализаторы не претерпевают изменений и работают десятки лет.

Термическая очистка — это процесс окисления компонентов выбросов при высокой температуре [9, 11].

Следующий вид среды, который подвергается загрязнению нефтегазовыми объекта-

Таблица 6: Установки, применяемые для очистки атмосферного воздуха

Виды сред		
Атмосфера		
№	Методы очистки и применяемые аппараты	Описание (достоинства и недостатки)
1.	Сухая очистка (пылеуловители)	Процесс разделения загрязненного газа, при котором частицы пыли, находящиеся во взвешенном состоянии, отделяются под действием гравитационных и инерционных сил
2.	Электрическая очистка	Принцип действия заключен в ионизации газа. Электростатический заряд передается частицам пыли, которые двигаются к осадительным электродам, где скапливаются на их поверхности
3.	Фильтрация	Процесс разделения газа от пыли с использованием перегородки. Процесс осуществляется за счет разницы давлений по обе стороны перегородки. В качестве перегородки используют фильтры с гибкими пористыми перегородками; с полужесткими пористыми перегородками; с жесткими пористыми перегородками; с зернистыми перегородками
4.	Мокрая газоочистка	Процесс разделения газозвешей, при котором взвешенные частицы смачиваются жидкостью и отделяются от газа. Очистка происходит за счет действия сил инерции. Жидкостью является вода, водные растворы серной кислоты, соды или иные вещества

ми, это литосфера. Основные причины загрязнения:

- строительство, бурение и добыча скважины;
- применение буровых растворов;
- аварии на нефтегазовых объектах и т. д.

Все эти действия приводят к:

- отчуждению, осушению или подтоплению территории;
- загрязнению компонентов ОС взвешенными, химическими веществами;
- вырубке леса и изменению характера землепользования;
- гибели растительного покрова;
- шумовым, вибрационным и другим воздействиям.

Основными способами защиты, восстановления ОС от подобных воздействий являются:

- рекультивация нарушенных земель — комплекс мероприятий по восстановлению нарушенного почвенного покрова,

биоресурсов, природной и геологической среды;

- переход на замкнутый цикл водоснабжения буровой установки, что обеспечит снижение норм водопотребления и утилизацию БСВ;
- шламовые амбары — очаг для прилегающих территорий, где хранятся массы отходов бурения и т. д.;
- использование сорбентов (торф, мох, опилки, сено, полиуретан, целлюлоза и т. д.);
- биотехнологии, основанные на окислении нефтепродуктов микроорганизмами, в результате которого происходит расщепление углеводов, их минерализация и последующая гумификация.

Данные способы восстановления и предотвращения загрязнения ОС являются наиболее эффективными и экономически целесообразными. Однако период восстановления почвенно-растительного покрова может достигать до 25 лет.

Литература

- [1] Башкин В.Н. Экологические риски. Расчет, управление, страхование: учебное пособие. — М.: Высш. шк., 2007. — 360 с.
- [2] Белов С.В. Безопасность жизнедеятельности и защита окружающей среды (техносферная безопасность): учебник. — М.: Юрайт, 2010. — 672 с.
- [3] Об охране окружающей среды (с изменениями на 31 декабря 2017 года) (редакция, действующая с 1 января 2018 года): [Федеральный закон № 7-ФЗ от 10.01.2002 (ред. от 01.01.2018): принят Государственной Думой Российской Федерации 20 декабря 2001 года]. — М., 2002. — 22 с.
- [4] Тетельмин В.В., Язев В.А. Защита окружающей среды в нефтегазовом комплексе: учебник. — М.: Интеллект, 2013. — 352 с.
- [5] Голик В.И. Экологические, экономические и правовые аспекты разработки месторождений полезных ископаемых: учебное пособие. — М.: ИЦ РГУ нефти и газа, 2012. — 282 с.
- [6] Ягафарова Г.Г. Инженерная экология в нефтегазовом комплексе: учебное пособие. — Уфа: УГНТУ, 2007. — 334 с.
- [7] Белов С.В. Безопасность жизнедеятельности и защита окружающей среды (техносферная безопасность): учебник для бакалавров. — М.: Юрайт, 2014. — 52 с.
- [8] Бухгалтер Э.Б. Экология нефтегазового комплекса [Text]: выставочные материалы / Бухгалтер Э.Б., Голубева И.А., Лыков О.П. и др. — [S. l. : s. n.], 2003. — 415 с.
- [9] Ветошкин А.Г. Процессы и аппараты защиты окружающей среды: учебник. — М.: Высшая школа, 2008. — 640 с.
- [10] Москаленко А.П. Экономика природопользования и ресурсосбережения: учебник. — М.: Феникс, 2014. — 479 с.
- [11] Муравых А.И. Управление экологической безопасностью: учеб. пособие. М.: ОП ИЦ РАГС. — 286 с.
- [12] Кукин П.П., Колесников Е.Ю. Оценка воздействия на окружающую среду. Экспертиза безопасности: учебник. — М.: Юрайт, 2016. — 454 с.
- [13] Оздоева, А.Х. Методические подходы экологической оценки инвестиционной привлекательности проектов полезного использования попутного нефтяного газа / А.Х. Оздоева // Физико-техническая информатика. СРТ16-17: труды Международной научной конференции. — Москва-Протвино: Изд. ИФТИ, 2017. — С. 281-284.
- [14] Бондарева Т.И. Утилизация диоксида углерода в промышленности / Т.И. Бондарева, Ю.Г. Пикуллин // Экология и промышленность России. — 2003. — Январь. — С. 38-40.
- [15] Шакеров Р. Об экологической безопасности // ЭЖЮРИСТ. -2017. -№02 (953).
- [16] Формирование инновационно-промышленных кластеров в России: Монография / И.С. Тремасова, В.П. Скобелина — Германия: Palmarium Academic Publishing, 2012. 80 с.
- [17] Мазлова Е.А. Экологический менеджмент нефтегазовых предприятий: учебное пособие. — М.: ИЦ РГУ нефти и газа, 2014. — 199 с.
- [18] Алексеева Н.А., Ибрагимова А.В. Выбор инвестиционной стратегии инвестора при выборе оптимального проекта утилизации попутного нефтяного газа в условиях неопределенности. / Вестник Удмуртского университета -2014 г. № 1-с. 7-11.

Исследование зависимости результатов оценок радиационного риска АЭС с реактором типа ВВЭР от состава населения, проживающего вокруг АЭС (на примере Ростовской и Калининской АЭС)*

Д. А. Маринина¹, М. А. Берберова²

mardasha0703@gmail.com, maria.berberova@gmail.com

¹Московский физико-технический институт (государственный университет)

Москва, ²Международный центр по ядерной безопасности

Атомные электростанции, являясь сложными технологическими объектами, представляют собой источник повышенного риска, в частности, специфического риска радиационного воздействия. Получение количественных оценок радиационного риска является критичным для проведения мероприятий по снижению риска и предотвращению последствий аварий. Существующие методы оценки радиационного риска не учитывают влияние внешних факторов, таких как состав населения, географические особенности, антропогенное изменение среды и т. д. Результатом анализа риска является оценка физических и экономических показателей для Калининской АЭС, учитывающих возрастную состав населения, и сравнение полученных оценок с результатами для Ростовской АЭС.

Ключевые слова: Оценка риска, радиационное воздействие, население, возрастной состав, безопасность, атомная электростанция, ВВЭР.

Radiation risk assessment on the population living near NPP with VVER-type reactor and its dependence on the age composition*

D. A. Marinina¹, M. A. Berberova²

¹Moscow Institute of Physics and Technology

Moscow, ²International nuclear safety center

Nuclear power stations as a complex technological objects represent a source of increased risk, in particular, specific risk of radiation exposure. Obtaining quantitative estimates of radiation risk is critical for carrying out activities to reduce risk and prevent the consequences of accidents. Existing methods for assessing radiation risk do not take into account the influence of external factors, such as population composition, geographical features, anthropogenic change, etc. The result of the risk analysis is an assessment of physical and economic indicators for Kalinin NPP, taking into account the age composition of the population, and comparing obtained estimates with the results for Rostov NPP.

Keywords: Risk assessment, radiation impact, population, age composition, safety, nuclear power plant, VVER.

Введение

Методологии расчета показателей риска таких сложных систем, как атомные электростанции, начали формироваться сравнительно недавно. Успешность их применения на данный момент существенно зависит от выполнения множества критериев, таких как доступность релевантных данных о поведении системы в многочисленных состояниях, существование схожих объектов, наличие данных о внешних факторах и т. д. Перечень мероприятий по управлению безопасностью АЭС напрямую зависит от оценок, полученных на основании данных рас-

четов, поэтому критически важным является усовершенствование методов оценки радиационного риска, как имеющего наиболее серьезные последствия.

Актуальность оценки риска радиационного воздействия на население с учетом возрастного критерия, как наиболее значимого, была сформулирована авторами ранее в работе, посвященной корректировке оценок радиационного риска для Ростовской АЭС [1]. Доступность данных о возрастном составе позволяет провести точные оценки и обосновать значимость возрастного критерия при оценке радиационного риска населения.

В 2012 году был выполнен отчет «Разработка перечня мероприятий по управлению

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 16-07-00089, 18-07-20030, 18-37-10008

безопасностью и оценка показателей риска для 2 очереди Калининской АЭС с реактором ВВЭР-1000 (В-320)» [2] в соответствии с «Методикой оценки показателей риска АЭС, адаптированной к АЭС с ВВЭР» [3]. В отчете в качестве примера АЭС рассматривается Калининская АЭС с реактором типа ВВЭР-1000.

В работе скорректирована оценка радиационного воздействия на население, проведенная в [2, 4], с учетом возрастного состава населения. Также авторами был проведен сравнительный анализ результатов, полученных в работе [1] для Ростовской АЭС, и результатов оценок риска для Калининской АЭС.

Краткое описание Калининской АЭС

Калининская атомная электростанция (рис. 1) расположена в Тверской области в 120 км от города Тверь на южном берегу озера Удомля и около одноименного города. Город Удомля расположен в северной части Тверской области, где проходит водораздел между Балтийским и Каспийским морями, между водосборными площадями рек Волги и Невы.



Рис. 1: Калининская АЭС. Общий вид

Первый энергоблок Калининской АЭС введен в эксплуатацию 9 мая 1984 года, второй – 11 декабря 1986 года. В 2004 и 2011 годах были запущены третий и четвертый энергоблоки.

Калининская АЭС занимает существенное место в экономике региона, вырабатывая

70 % электроэнергии, производимой в Тверской области.

Положение площадки Калининской АЭС в географическом плане определяется следующими координатами: широта $57^{\circ}54'$, долгота $35^{\circ}04'$, абсолютная отметка планировки 161,0 м [2].

В административном отношении площадка АЭС расположена в Удомельском, районе Тверской области, в 7 км к северо-востоку от г. Удомля. Расстояние до областного центра – г. Тверь составляет 115 км по прямой [2].

30-км зона Калининской АЭС частично захватывает территории 4-х административных районов Тверской области – Удомельского, Вышневолоцкого, Максатихинского и Лесного [2].

Кратчайшие расстояния до ближайших административных границ соседних областей составляют [2]:

- Новгородская область – 25 км к северо-западу;
- Вологодская область – 105 км к северо-востоку;
- Ярославская область – 135 км к востоку;
- Московская область – 150 км к юго-востоку;
- Смоленская область – 205 км к югу;
- Псковская область – 280 км к юго-западу.

Тверская область не граничит с государствами ближнего и дальнего зарубежья [2].

Площадка АЭС расположена на южном берегу оз. Удомля, в 2,7 км восточнее оз. Песьво. Озера соединены прорезью и используются для технического водоснабжения и охлаждения систем АЭС. Сток воды в р. Съезжа из озер Удомля и Песьво зарегулирован с 1987 г. гидроузлом с НПП-156,25 м. абс [2].

Другие крупные озера Кубыча, Наволок и Кезадра находятся соответственно в 6,0 км к юго-западу, 13,5 км к северу и 15,0 км

к северо-востоку от площадки АЭС. Район также характеризуется наличием других более мелких озер, таких как Саминец, Мишутинское, Бол. Кжемля. Наиболее крупными реками района являются Волчина, протекающая юго-западнее площадки АЭС, Съежа и Тихомандрица, расположенные к северо-западу от АЭС.

Площадка АЭС связана с ближайшим крупным транспортным узлом г. Удомлей железнодорожным и автомобильным путями. На расстоянии 3 км южнее площадки проходит железнодорожная магистраль общего пользования Бологое-Сонково. Ближайшие крупные населенные пункты находятся от площадки на расстоянии соответственно Вышний Волочек — 48 км и Бологое — 57 км. Из других ближайших городов Тверской области следует отметить Торжок (86 км), Бежецк (95 км), Осташков (138 км), Кимры (166 км), а также г. Боровичи Новгородской области, расположенный в 80 км от площадки АЭС. Жилой поселок АЭС примыкает к южной стороне существующей застройки г. Удомля и связан с промплощадкой автомобильной дорогой с твердым покрытием [2].

На расстоянии 22 км к югу от площадки АЭС проходит федеральная автомобильная дорога Вышний Волочек — Бежецк.

В санитарно-защитной зоне Калининской АЭС населенные пункты отсутствуют [2].

В ближайшем окружении от АЭС (в 30-км зоне) отсутствуют промышленные или какие-либо другие предприятия (склады, шахты, карьеры, буровые установки и др.), способные повлиять на работу станции. Магистральных газо- и нефтепроводов в районе площадки АЭС нет. Газопроводов в поселке АЭС нет, в жилых домах — электроплиты. В старой застройке г. Удомля используется баллонный газ [2].

Артскважины питьевого водоснабжения АЭС и станция обезжелезивания удалены от промплощадки на расстояние 6 км и размещаются за жилпоселком в южном направлении от АЭС [2].

Промышленность 30-км зоны станции представлена предприятиями Максатиха находится в населенном пункте с одноименным названием на расстоянии 50 км к юго-востоку от площадки АЭС [2].

30-км зона АЭС включает в себя территории охотничьих хозяйств, расположенных в 15 км к северо-востоку от АЭС в районе оз. Кезадра, в 21 км к северо-востоку от промплощадки в районе населенных пунктов Лишутино, Гирино, Мал. Званица, Свирка, а также в 15 км к югу от АЭС в районе озер Молдино, Перхово, Судомля, Волчино, Олынево [2].

Территория радиусом 30 км от границ площадки Калининской АЭС географически, в административном и хозяйственном отношении входит в состав Тверской области, располагаясь в северо-западной ее части, в пределах Удомельского, Вышневолоцкого, Максатихинского и Лесного муниципальных районов. Расположение Калининской АЭС в административном отношении приведено на рис. 2 [2].



Рис. 2: Административная карта района Калининской АЭС, R=30 км

Современное размещение населения обусловлено, в первую очередь, наличием промышленных и сельскохозяйственных пред-

приятый, а также наличием дорог и объектов инфраструктуры [2].

По данным о численности населения на 1 января 2008 г., крупными населенными пунктами с числом жителей более 300 чел. на рассматриваемой территории являются: город Удомля с населением 31803 чел., п. Брусово (836 чел.), д. Ряд (562 чел.), с. Молдино (458 чел.), с. Котлован (351 чел.), д. Дятлово (452 чел.), п. Овсище (323 чел.) [2].

Все эти населенные пункты привязаны к транспортным магистралям. Вдоль северной границы территории прослеживаются места, где население отсутствует вообще, это так называемые районы демографического бедствия: район озера Кезадра, местность севернее д. Гриблянка [2].

На 10-км территории показатели численности и плотности населения определяет г. Удомля, расположенный в юго-западном направлении от АЭС (рис. 3) [2].



Рис. 3: План г. Удомля

Численность населения Тверской области на 01.01.08 составляет 1379,5 тыс. чел. [2].

Доля 30-км территории Калининской АЭС в общей численности населения области составляет 3,2% [2].

На территории радиусом 30 км от границ площадки Калининской АЭС проживает 44,49 тыс. чел., в том числе [2]:

- в городе Удомля – 32 тыс. чел. или 71,5% всего населения рассматриваемой территории;
- в 19 сельских поселениях (323 населенных пункта) – 12,7 тыс. чел., или 28,5% всего населения 30-км территории.

93% населения 30-км территории, что составляет 41,4 тыс. чел., проживают в населенных пунктах Удомельского муниципального района. Из них городское население составляет 76,8%, сосредоточенное в г. Удомля (31,8 тыс. чел.), 23,2% — сельское население, проживающее в деревнях и селах, расположенных на территориях Брусовского, Еремковского, Зареченского, Копачевского, Котлованского, Куровского, Молдинского, Порожжинского, Рядского, Таракинского, Удомельского сельских поселений (9,6 тыс. чел.) [2].

Население Вышневолоцкого муниципального района, проживающее на территории 30 км от Калининской АЭС, составляет 2221 чел. (5% всего населения территории) и рассредоточено в 73 населенных пунктах, относящихся к Овсищенскому, Дятловскому и Сорокинскому сельским поселениям. Наиболее крупными населенными пунктами являются: п. Овсище с численностью 323 чел. — центр Овсищенского сельского поселения, д. Дятлово (452 чел.) — центр Дятловского сельского поселения, с. Осечно (207 чел.), д. Кузнецово (227 чел.), д. Старое (140 чел.) В остальных населенных пунктах, являющихся деревнями, проживает не более 100 чел. в каждом. В семнадцати населенных пунктах на 01.01.08 нет постоянно проживающего населения [2].

В 30-км территорию входят 3 сельских поселения Максатихинского муниципального района – Каменское, Труженицкое и Малышевское, общей численностью 875 чел. или 2% всего населения территории. Количество населенных пунктов – 26, из них 23 деревни и 3 хутора. Наиболее крупные на-

селенные пункты – д. Каменка насчитывает 293 жителей, д. Ново-Пхово насчитывает 147 жителей, д. Починок 126 жителей, численность остальных населенных пунктов составляет менее 100 чел. в каждом. В 7 населенных пунктах, попадающих на 30-км территорию, нет постоянно проживающего населения [2].

В 30-км территорию Калининской АЭС также входит часть территории Лесного муниципального района, на которой расположена 1 деревня, относящаяся к Соргожскому сельскому поселению, в которой на 01.01.08 отсутствует постоянно проживающее население, и город Удомля, расположенный на расстоянии 3 км от площадки АЭС и насчитывающий на 01.01.08 31,8 тыс. чел. Остальные населенные пункты, попадающие на рассматриваемую территорию, сельского типа [2].

Расстояние от областного центра г. Твери до площадки Калининской АЭС составляет 115 км по прямой. Из других ближайших городов Тверской области следует отметить г. Боровичи Новгородской области с населением 55 тыс. чел., расположенный в 79 км от АЭС [2].

Города с населением 100 тыс. чел. и более на территории в радиусе 100 км от АЭС отсутствуют [2].

Критический населенный пункт территории 30-км зоны Калининской АЭС – г. Удомля, критическая группа населения (по газоаэрозольному выбросу АЭС) – население ближайших к станции окраин города [2].

Методология оценки радиационного риска Калининской АЭС

Оценка риска от чрезвычайной ситуации (ЧС) на Калининской АЭС [2] выполняется по упрощенной процедуре проведения вероятностного анализа безопасности (ВАБ) с помощью компьютерной программы FOOD (версия 2), которая позволяет рассчитать дозы облучения населения на различных

расстояниях от источника выброса радионуклидов.

На основании ранее проведенного ВАБ в качестве наиболее вероятной аварии на Калининской АЭС выбрана авария «Средняя течь первого контура» (с частотой реализации $4.2 \cdot 10^{-4}$ 1/реактор-год), в качестве наиболее опасной аварии – «Большая течь первого контура с отказом систем планового и аварийного расхолаживания и аварийного ввода бора низкого давления» (с частотой $1.47 \cdot 10^{-6}$ 1/реактор-год).

Далее будет рассматриваться только наиболее опасный сценарий ЧС. Радиационные последствия аварии выполняются в предположении, что высота выброса над уровнем земли 40 м, продолжительность выброса 24 часа. Выброс радионуклидов в окружающую среду представлен в таблице 1.

Таблица 1: Выброс радионуклидов в окружающую среду при наиболее опасной аварии «Большая течь первого контура с отказом систем планового и аварийного расхолаживания и аварийного ввода бора низкого давления» на Калининской АЭС

Нуклид	Количество продуктов деления, Ки
131I	$1.35 \cdot 10^5$
133I	$5.7 \cdot 10^4$
135I	$2.16 \cdot 10^4$
85mKr	$7.2 \cdot 10^4$
87Kr	$4.5 \cdot 10^4$
88Kr	$1.2 \cdot 10^5$
133Xe	$1.4 \cdot 10^7$
135Xe	$3.0 \cdot 10^5$
134Cs	$2.4 \cdot 10^2$
137Cs	$1.5 \cdot 10^2$

Выброс радиоактивных веществ в окружающую среду считается кратковременным. При оценке доз облучения населения учитываются внешнее облучение (от радиоактивного облака и загрязненной поверхности земли) и внутреннее облучение (прямая ингаляция).

Физическими показателями радиационных последствий аварии на население являются:

- поглощенная доза внешнего облучения всего тела и внутреннего облучения щитовидной железы при ингаляции на начальном этапе аварии (первые 10 суток);
- годовая эффективная доза облучения.

В [2] оценка производится для среднестатистического взрослого человека со средней радиочувствительностью. В таблице 2 приведены результаты оценки поглощенных доз облучения взрослого населения.

Таблица 2: Оценка поглощенных доз облучения взрослого населения на начальном этапе (первые 10 суток) при наиболее опасной аварии на Калининской АЭС

Расстояние от АЭС, км	Поглощенная доза внешнего облучения на оси распространения, мГр	Поглощенная доза внутреннего облучения на оси распространения, мГр
3	14	480
5	5.9	200
10	1.9	63
15	0.9	30
20	0.56	18
30	0.37	12
40	0.28	9
50	0.11	3.4
100	0.03	0.87

Оценка доз облучения с учетом возрастного состава населения, проживающего вблизи Калининской АЭС

Согласно [5], оценка доз облучения населения при радиоактивном загрязнении проводится для трех возрастных групп:

- 1-я группа – для взрослого населения (> 18 лет);
- 2-я группа – для детей школьного возраста (7-12 лет);
- 3-я группа – для дошкольников (1-2 года).

Значение эффективной дозы внешнего облучения при погружении в радиоактивное облако и от подстилающей поверхности ли-

нейно зависит от возрастного коэффициента $K_{i,a}$ [5].

$K_{i,a}$ принимается равным:

- 0.7 мЗв/мГр для взрослого населения;
- 0.75 мЗв/мГр для детей школьного возраста;
- 0.85 мЗв/мГр – для дошкольников.

Доля поглощенного внутреннего облучения в суммарном поглощенном облучении составляет около 97% (таблица 2). Следуя методологии расчетов, описанной в [1], корректировка суммарной эффективной дозы E_i^{sum} осуществляется только на коэффициент для внутреннего облучения.

Значение ожидаемой эффективной дозы E_i^{thr} (мЗв) на щитовидную железу рассчитывается с помощью выражения:

$$E_i^{thr} = V_i \cdot T \sum C_a^k \cdot h_{i,thr}^k, \quad (1)$$

где: $h_{i,thr}^k$ – дозовый коэффициент для i -ой группы населения и k -ого радионуклида (теллур или йод), мЗв/кБк; C_a^k – средняя концентрация k -ого радионуклида в приземном слое воздуха за время прохождения радиоактивного облака, кБк/м³; V_i – интенсивность дыхания представителей i -ой группы населения, м³/час; T – время облучения от данного источника, час.

Интенсивность дыхания V_i для трех возрастных групп равна:

- 1.4 м³/час для взрослого населения;
- 1.1 м³/час для детей школьного возраста;
- 0.35 м³/час – для дошкольников.

Расчет средневзвешенного дозового коэффициента $h_{i,thr}^k$ для групп населения проведен на основании количества продуктов деления источников (таблица 1) и представлен в таблице 3.

Таблица 3: Дозовые коэффициенты для трех возрастных групп населения и изотопов йода)

Нуклид	$h_{3,thr}$, мЗв/кБк	$h_{2,thr}$, мЗв/кБк	$h_{1,thr}$, мЗв/кБк
131I	$1.40 \cdot 10^0$	$3.70 \cdot 10^{-1}$	$1.50 \cdot 10^{-1}$
133I	$3.50 \cdot 10^{-1}$	$7.40 \cdot 10^{-2}$	$2.80 \cdot 10^{-2}$
135I	$7.00 \cdot 10^{-2}$	$1.50 \cdot 10^{-2}$	$5.70 \cdot 10^{-3}$
Ср.зн.	$9.85 \cdot 10^{-1}$	$2.55 \cdot 10^{-1}$	$1.03 \cdot 10^{-1}$

Таблица 4: Оценка средних суммарных эффективных годовых доз внешнего и внутреннего облучения групп населения в каждом из кольцевых сегментов румба при наиболее опасной аварии

Расстояние от АЭС, км	E_1^{sum} , мЗв	E_2^{sum} , мЗв	E_3^{sum} , мЗв
3-5	33.00	64.31	79.03
5-10	12.95	25.24	31.01
10-15	4.40	8.58	10.54
15-20	2.35	4.58	5.63
20-30	1.50	2.92	3.59
30-40	1.03	2.00	2.45
40-50	0.59	1.15	1.41
50-100	0.21	0.40	0.49

Таблица 5: Распределение коллективной дозы для групп населения по румбам в 100 километровой зоне вокруг Калининской АЭС при наиболее опасной аварии

Румб	N_1	N_2	N_3	E_1^{col}	E_2^{col}	E_3^{col}
С	26378	2350	2088	10.86	1.87	2.03
СВ	26111	2326	2067	9.36	1.61	1.75
В	54382	4844	4306	19.30	3.31	3.61
ЮВ	37230	3316	2948	15.75	2.71	2.95
Ю	74937	6675	5933	261.46	45.33	49.51
ЮЗ	82931	7387	6566	494.01	85.74	93.64
З	54136	4822	4286	224.53	38.94	42.53
СЗ	73322	6531	5805	20.70	3.54	3.85
Итого	429427	38249	33999	1056	183	200

В выражении (1) два коэффициента зависят от возрастного состава населения. Следовательно, перемножение $h_{i,thr}^k$ и V_i дает совокупный коэффициент, показывающий линейную зависимость эффективной дозы на щитовидную железу от возраста облучаемого населения:

- 0.1440 ($\text{м}^3 \cdot \text{мЗв}$)/($\text{час} \cdot \text{кБк}$) для взрослого населения;
- 0.2806 ($\text{м}^3 \cdot \text{мЗв}$)/($\text{час} \cdot \text{кБк}$) для детей школьного возраста;
- 0.3448 ($\text{м}^3 \cdot \text{мЗв}$)/($\text{час} \cdot \text{кБк}$) — для дошкольников.

В таблице 4 приведены результаты расчета суммарной эффективной дозы E_i^{sum} с учетом корректировки на коэффициент $h_{i,thr}^k \cdot V_i$ для внутреннего облучения.

Для оценки радиационного воздействия выбросов и планирования аварийных мероприятий использовались количественные показатели зависимости численности населения от расстояния от АЭС. Плотность населения вне 30-км зоны принята средней для региона Калининской АЭС. Для оценки количества населения N_i в каждой возрастной группе, использовались данные Всероссийской переписи населения 2010 года по Тверской области, согласно которой доля взрослого населения составляет 85%, школьников (7-12 лет) — 8%, дошкольников (1-2 года) — 7%.

Для оценки радиационных последствий аварии на различные группы населения в 8 румбах в 100 километровой зоне вокруг КЛнАЭС выполнены расчеты:

- коллективной дозы облучения E_i^{col} , чел. · Зв/год;
- количества поздних смертей, N_i^{dth} ;
- коллективного риска поздней смерти, R_i^{dth} .

Риск ранних смертей среди населения исключается, т.к. на начальном этапе аварии предполагаемые индивидуальные поглощенные дозы внешнего облучения всего тела не достигают уровня 1 Гр, внутреннего облучения щитовидной железы при ингаляции не достигают уровня 5 Гр, при которых возможны ранние детерминированные эффекты [2, 4].

Экономический ущерб G (млн. руб) в результате причинения вреда здоровью населения, который может проявиться в виде стохастических эффектов (радиогенные

Таблица 6: Распределение риска поздней смерти для групп населения по румбам в 100 километровой зоне вокруг Калининской АЭС при наиболее опасной аварии

Румб	N_1^{dth}	N_2^{dth}	N_3^{dth}	R_1^{dth}	R_2^{dth}	R_3^{dth}
С	0.46	0.11	0.12	$6.7 \cdot 10^{-7}$	$1.5 \cdot 10^{-7}$	$1.7 \cdot 10^{-7}$
СВ	0.39	0.09	0.10	$5.8 \cdot 10^{-7}$	$1.3 \cdot 10^{-7}$	$1.5 \cdot 10^{-7}$
В	0.81	0.19	0.21	$1.2 \cdot 10^{-6}$	$2.8 \cdot 10^{-7}$	$3.0 \cdot 10^{-7}$
ЮВ	0.66	0.15	0.17	$9.7 \cdot 10^{-7}$	$2.3 \cdot 10^{-7}$	$2.5 \cdot 10^{-7}$
Ю	10.98	2.58	2.82	$1.6 \cdot 10^{-5}$	$3.8 \cdot 10^{-6}$	$4.1 \cdot 10^{-6}$
ЮЗ	20.75	4.89	5.34	$3.0 \cdot 10^{-5}$	$7.2 \cdot 10^{-6}$	$7.8 \cdot 10^{-6}$
З	9.43	2.22	2.42	$1.4 \cdot 10^{-5}$	$3.3 \cdot 10^{-6}$	$3.6 \cdot 10^{-6}$
СЗ	0.87	0.20	0.22	$1.3 \cdot 10^{-6}$	$2.9 \cdot 10^{-7}$	$3.2 \cdot 10^{-7}$
Итого	44.35	10.43	11.39	$6.5 \cdot 10^{-5}$	$1.5 \cdot 10^{-5}$	$1.7 \cdot 10^{-5}$

рак и наследственные заболевания), оценивается по формуле [2]:

$$G = z \cdot E^{col}, \quad (2)$$

где: z – денежный эквивалент потери одного человеко-года жизни населения принимается равным 0,3 млн.руб/чел-Зв; E^{col} – коллективная (популяционная) эффективная доза населения, чел-Зв.

Для обоснования расходов на радиационную защиту при реализации принципа оптимизации принимается, что облучение в коллективной эффективной дозе в 1 чел-Зв приводит к потенциальному ущербу, равному потере примерно 1 чел-года жизни населения [2].

В таблице 7 приведено сравнение полученных оценок физических и экономических показателей радиационного риска для населения, проживающего вокруг Калининской АЭС, при наиболее опасной аварии с оценками, приведенными в [2].

В таблице 8 приведено сравнение показателей, полученных для Калининской АЭС, с показателями, рассчитанными авторами для Ростовской АЭС [1]

Выводы

По результатам выполненных оценок доз и последствий облучения населения при рассмотренной аварии на Калининской АЭС можно сделать следующие выводы:

1. Корректировка на возрастной состав дает более высокие показатели риска облучения населения.
2. Аварийное облучение населения с учетом корректировки находится ниже уровней, на которые распространяются требования радиационной безопасности [6].
3. Необходимо учитывать приведенные оценки при разработке мер по снижению риска и смягчения последствий для наиболее чувствительных возрастных групп населения (1-12 лет).

Заключение

В работе проведено исследование зависимости результатов оценок радиационного риска АЭС с реактором типа ВВЭР от состава населения, проживающего вокруг АЭС (на примере Ростовской и Калининской АЭС). Проведено сравнение полученных оценок физических и экономических показателей радиационного риска для населения, проживающего вокруг Калининской АЭС, при наиболее опасной аварии с ранее выполненными расчетами для Ростовской АЭС. Доказана необходимость учета проведенных оценок при разработке мер по снижению риска и смягчения последствий для наиболее чувствительных возрастных групп населения (1-12 лет).

Таблица 7: Сравнение полученных оценок физических и экономических показателей радиационного риска для населения, проживающего вокруг Калининской АЭС, при наиболее опасной аварии

Возрастной состав	E^{col}	N^{dth}	R^{dth}	G
Без учета возрастного состава	1230	61.3	$1.48 \cdot 10^{-5}$	369
С учетом возрастного состава	1439	66.2	$9.73 \cdot 10^{-5}$	431.7
Отклонение	17%	8%	557%	17%

новлением Главного государственного санитарного врача РФ // No.47 от 07.07.2009. — М., 2009. — 7 с.

Таблица 8: Сравнение показателей, полученных для Калининской АЭС, с показателями, рассчитанными для Ростовской АЭС

Возрастной состав	E^{col}	N^{dth}	R^{dth}	G
Без учета возрастного состава	234	11.7	$5.50 \cdot 10^{-8}$	70.2
С учетом возрастного состава	279	12.9	$28.20 \cdot 10^{-8}$	83.7
Отклонение	19%	10%	413%	19%

Литература

- [1] Маринина Д.А., Берберова М.А. «Оценка риска радиационного воздействия на население, проживающего вблизи рассматриваемой АЭС с реактором типа ВВЭР, с учетом возрастного состава» // Труды Международной научной конференции «Ситуационные центры и ИАС4i для мониторинга и безопасности» SCVRT2017, 28 ноября 2017 г. Протвино-Москва, 2017: Изд. ИФТИ. — С. 94-98.
- [2] Жуков И.В., Берберова М.А., и др. Разработка перечня мероприятий по управлению безопасностью и оценка показателей риска для 2 очереди Калининской АЭС с реактором ВВЭР-1000 (В-320): отчет о НИР рег. № 2339ОТ12 // Международный Центр по Ядерной Безопасности — М., 2012. — 109 с.
- [3] Разработка проекта Методики оценки показателей риска АЭС, ее согласование с соисполнителями и представление на утверждение в установленном порядке: отчет о НИР рег. № 1347АО08 // Международный Центр по Ядерной Безопасности — М., 2008. — 132 с.
- [4] Берберова М.А. Оценка риска для АЭС // Атомная энергия, т. 109, вып. 6. — М., 2010. — С. 307-311.
- [5] Оперативная оценка доз облучения населения при радиоактивном загрязнении территории воздушным путем: [Методические указания МУ 2.6.1.2153-06: утверждены Главным Государственным санитарным врачом РФ 04.12.2006]. — М., 2007. — 90 с.
- [6] СанПиН 2.6.1.2523-09. Нормы радиационной безопасности (НРБ-99/2009). Утверждены Поста-

Геотектоника, как причина различия результатов сравнительной оценки внутреннего и внешнего облучения и ущерба населению от АЭС*

Л. Х. Оздоева
luisana54@bk.ru

Московский Государственный Университет имени М. В. Ломоносова

В данной работе проводится анализ различий в тектоническом строении и развитии рельефа АЭС и их влияние на результат оценок риска АЭС (на примере Ростовской и Калининской АЭС). Ключевые слова: геотектоника, сравнительная оценка, внутреннее и внешнее облучение, население, АЭС.

Ключевые слова: геотектоника, сравнительная оценка, внутреннее и внешнее облучение, население, АЭС.

Geotectonics, as a reason for the difference in the results of a comparative assessment of internal and external exposure and damage to the population from nuclear power plants*

L. Kh. Ozdоеva
Moscow State University

In this paper, we analyze the differences in the tectonic structure and the development of the topography of nuclear power plants and their effect on the result of NPP risk assessments (for example, the Rostov and Kalinin NPP).

Keywords: geotectonics, comparative evaluation, internal and external exposure, population, NPP.

Введение

В пределах Восточно-Европейского региона, наиболее экономически развитой части России, построены десять гидравлических, тепловых и атомных электростанций. Они возведены в разных инженерно-геологических условиях, на отложениях различных формаций и генетических типов. В исследовании проведем анализ обобщение инженерно-геологических условий Калининская АЭС и Ростовская АЭС. Рассмотрим ряд факторов, влияющих на их строительство, функционирование.

Мощность толщи четвертичных отложений. Если она сравнительно невелика (до 30 м), то наиболее ответственные сооружения можно возводить непосредственно на скальных или полускальных породах либо использовать для передачи давления на них свайные основания. Если мощность четвертичных отложений больше, такие решения становятся технически слишком сложными и дорогостоящими и приходится использо-

вать в качестве основания дисперсные отложения, прибегая в нужных случаях к их упрочнению методами технической мелиорации.

Коренные породы, используемые в качестве оснований атомных станций, относятся в основном к метаморфическим, изверженным или осадочным — терригенным и карбонатным. Проблемы, возникающие при возведении таких объектов на перечисленных породах, существенно различны: в первом случае основания вполне надежны по своим прочностным и деформационным свойствам, но существенные трудности нередко создает разрывная тектоника, характерная для кристаллического фундамента платформы: терригенные породы менее надежны и, кроме того, нередко как разуплотнение, размокание и склонны к таким процессам выветривание; при строительстве на карбонатных породах главной проблемой оказывается карст и карстово-суффозионные процессы.

При большой мощности четвертичных отложений целесообразно использовать для классификации оснований общепринятый инженерной геологии генетический прин-

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 16-07-00089, 18-07-20030, 18-37-10008

цип, поскольку различные по генезису отложения заметно различаются по строению, условиям залегания, литологическому составу и, что особенно важно, инженерно-геологическим свойствам. По этому признаку выделены площадки на ледниковых аллювиальных и лессовых отложениях. Кроме того, в отдельный тип выделены площадки, сложенные глинистыми породами разного генезиса.

Инженерно-геологические условия АЭС

Калининская и Ростовская АЭС находятся на Европейской платформе. Основания, сложенные четвертичными отложениями при глубоком залегании скалистых пород. Мощные толщи четвертичных отложений на территориях Калининской и Ростовской АЭС различаются генезисом, составом, состоянием и физико-механическими свойствами. Калининская АЭС расположена в области ледниковой аккумуляции; Ростовская АЭС в районах распространения лёссовых грунтов.

Важнейшая особенность площадок этого типа — значительная неоднородность строения ледниковой толщи с характерным переслаиванием пород разного состава, измене мощности. Большой изменчивостью отличается и общая мощность ледниковой толщи, так как она обычно подстилается сильно зродированным основанием. Изменчивость строения толщ, прочностных, деформационных и фильтрационных свойств пород создает опасность неравномерных осадок. Как правило значительной сложностью характеризуются также гидрогеологические условия. Влияние их на строительство и эксплуатацию сооружений многообразно: усложнение технологии и организации работ, ухудшение свойств массива грунтов в процессе техногенного изменения гидрогеологических условий, активизация или проявление неблагоприятных геологических процессов, ухудшение экологической обстановки в связи с загрязнением подземных вод и т. д. Отмеченные особенности областей ледниковой аккумуля-

ции нередко создают большие трудности при разработке и обосновании рациональных проектных решений и их практическом осуществлении. Показательна в этом отношении Калининская АЭС [1].

Калининская АЭС

Калининская АЭС расположена на берегу озера Удомля, используемое для оборотной системы охлаждения станции. Как видно из разреза (рис. 1), в строении ледниковых отложений принимают участие две морены и флювиогляциальные отложения, подстилаемые мощной толщей известняков [2].

Участок размещения станции находится на границе Московской синеклизы и Балтийского щита, разделенных Валдайским авлакогеном, который до известной степени определяет сейсмичность площадки: проектное землетрясение (ПЗ) 5 баллов, максимальное расчетное землетрясение (МРЗ) 6 баллов. Интенсивность современных тектонических движений близка к нулю. Гидрогеологические условия характеризуются наличием подземных вод в ледниковых отложениях и известняках карбона.

Сложность геологических и гидрогеологических условий площадки не была в достаточной мере учтена при проектировании станции и ведении строительных работ. Это вызвало ряд осложнений при возведении первого энергоблока и потребовало корректировки проектных решений при строительстве второй очереди станции. В частности, наблюдались значительные осадки и крен фундамента реакторного отделения первой очереди: причинами явились длительное нахождение котлована реакторного отделения открытым, что вызвало неравномерное выветривание и разуплотнение глинистых грунтов за счет фильтрационных сил и морозного пучения; подрезка участка размещения реакторного отделения котлованом главного корпуса; длительное и неупорядоченное водопонижение при строительстве около озера блочных насосных станций (БНС) с заложением дна котлована 16 м. Большой дебит откачки (220

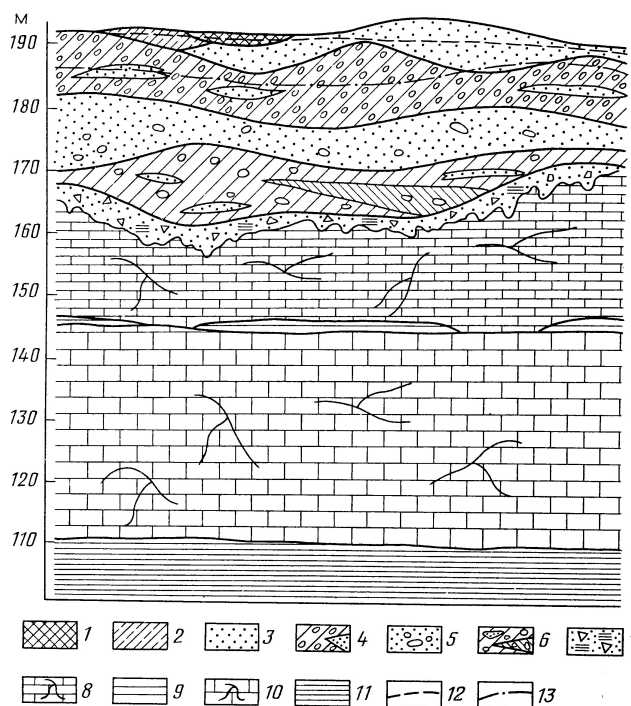


Рис. 1: Схематический геологический разрез площадки Калининской АЭС: 1 – насыпные грунты; 2 – верхнечетвертичные озерно-ледниковые глины; 3 – валдайские надморенные флювиогляциальные отложения (пески разной крупности); 4 – валдайская морена (валунные суглинки и глины с линзами песков); 5 – московско-валдайские флювиогляциальные отложения (пески с гравием, галькой и валунами); 6 – московская морена (валуны суглинки с линзами песков); 7 – элювий известняков (дресвяно-щебнистая масса, переходящая в известковистый алевроит и песок); 8 – подольский горизонт (известняки частично доломитизированные); 9, 10 – каширский горизонт (9 – глины мергелистые, 10 – известняки трещиноватые); 11 – верейский горизонт (глина); 12 – уровень грунтовых вод; 13 – пьезометрический уровень напорных вод

тыс. м³/сут), значительный радиус влияния (до 3 км), существенная срезка напора в четвертичном и каменноугольном водоносных горизонтах и многочисленные остановки насосов вызвали, по образному выражению геологов Горьковского отделения Атомэнергопроекта, «раскачку» ледниковых отложений, что способствовало разрушению их структурных связей, развитию объемных деформаций грунтов и, возможно, проявлению суффозионных процессов. Характерно, что при строительстве одной из станций, удаленной от РО примерно на

500 м, осадки реакторного отделения возобновились.

Отмеченный комплекс неблагоприятных техногенных воздействий на геологическую среду привел не только к значительным осадкам реакторного отделения (средняя осадка в строительный период составила 27 см, что в 2,5 больше расчетной), но и к развитию крена (13 см), составляющего до 50% осадки. Это потребовало проведения в процессе строительства специальных мероприятий, в частности пригрузки с помощью укладки слоя бетона той части сооружения, где осадки «отставали». Темпы строительства БНС 2 были повышены (сроки водопонижения сокращены до 9 месяцев, перемены в откачке воды исключены); строительство БНС 3 и 4 выполнено в едином комплексе. Реакторное отделение и машинное отделение второй очереди станции закладывались на одной отметке [2].

Ростовская АЭС

Ростовская АЭС, расположенная на южном побережье Цимлянского водохранилища, в районах распространения лессовых грунтов. Основные сооружения возводились на свайном основании с прорезкой сваями всей толщи лессовых грунтов. При проектировании станции рассматривалось семь конкурирующих площадок. Предпочтение отдано площадке с небольшой мощностью лессовых грунтов, достаточно надежными подстилающими породами глубоким залеганием подземных вод. В геоморфологическом отношении участок строительства расположен на погребенной террасе Дона и прилегающем склоне Доно-Сальского водораздела в геоструктурном — на северо-западной границе Цимлянской зоны поднятий в 25 км от Донбасско-Астраханского разлома.

В геологическом строении площадки принимают участие лессы первого типа по просадочности мощностью 2-7 м, ергенинские пески мощностью 12-17 м и Майкопские глины. Сейсмичность площадки с учетом наблюдающегося и прогнозируемого обводнения и резонансного эффекта оценена сле-

дующим образом: проектное землетрясение (ПЗ) 6 баллов, максимальное расчетное землетрясение (МРЗ) 7 баллов. Гидрогеологические условия характеризуются наличием двух гидравлически связанных горизонтов: в четвертичных отложениях и ергенинских песках.

В первую очередь строительства Ростовской АЭС входило возведение четырех энергоблоков. Неровная кровля ергенинских песков обусловила различную мощность лессовидных суглинков. Это повлияло на выбор типов фундаментов для основных сооружений. Так, для первого, третьего и четвертого блоков были предусмотрены монолитные железобетонные ростверки на сваях длиной 12-14, 8-10 и 9-11 м, соответственно, для второго блока — фундаментная плита на естественном основании — ергенинских песках (рис. 2) средней плотности [2].

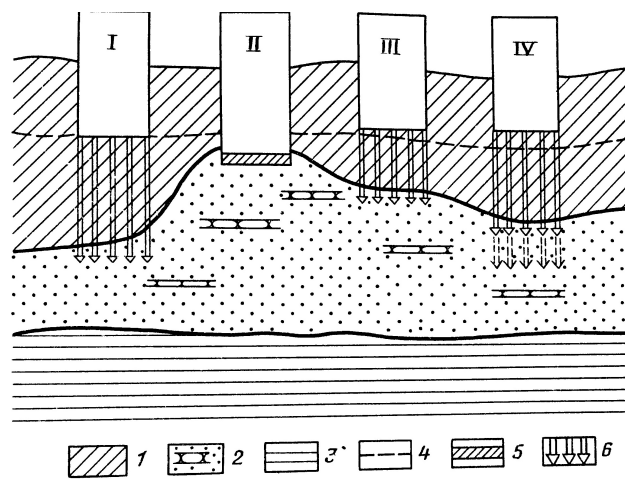


Рис. 2: Схематический геологический разрез площадки Ростовской АЭС с врезкой реакторных отделений: 1 — четвертичные лёссовые суглинки; 2 — неогеновые пески (ергенинская свита); 3 — палеоген-неогеновые глины; 4 — уровень подземных вод. I - IV — энергоблоки

Влияние на окружающую среду

По сравнению с традиционным видом энергетических установок — тепловыми станциями, работающим и на органическом топливе, атомные станции при эксплуатации не дают золы, шлаков, выбросов оксидов серы, азота и углерода. Радиационное загрязне-

ние окружающей среды также в несколько раз меньше, чем при эксплуатации тепловых станций. АЭС обладают определенными преимуществами и по сравнению с гидравлическими станциями, так как не вызывают затопления и подтопления обширных территорий. Таким образом, влияние атомных станций на окружающую среду меньше, чем энергетических объектов другого типа. Вместе с тем нельзя не учитывать, что в технологический процесс получения энергии АЭС входят такие звенья, как добыча, транспортировка и хранение атомного сырья, захоронение радиоактивных отходов, консервация или ликвидация объектов атомной энергетики, где создается значительный экологический риск, особенно в случае возникновения «нештатных» или аварийных ситуаций. В настоящее время многие из этих проблем еще не получили удовлетворительного решения и заслуживают всестороннего изучения.

Особенности взаимодействия атомных станций с геологической средой определяются следующими особенностями этих объектов:

1. Сложным характером застройки атомных станций, включающей здания и сооружения разного назначения (энергетические, гидротехнические, промышленные, транспортные, гражданские, специальные), оказывающие различное воздействие на геологическую среду, в результате чего формируется интегральная зона взаимодействия со сложной структурой, трудно поддающаяся изучению и оценке.
2. Многообразной природой оказываемого воздействия, сопровождающегося формированием гидродинамического теплового химического, электрического, радиационного, вибрационного и других полей, изменяющихся во времени и в пространстве и взаимодействующих как между собой, так и с геологической средой.
3. Жесткими допусками к различным видам воздействия на некоторые из основных сооружений.

Комплексное влияние различных видов техногенных воздействий АЭС на геологическую среду и взаимосвязь этих изменений показаны на рис. 3 [2]. Учитывая современные позиции инженерной геологии и требования рациональному использованию и охране природной среды, все изменения геологической среды объединены на схеме в две группы: 1) геоэкологические, приводящие к необратимым изменениям геосреды, нарушающим ее природное равновесие; 2) геотехнические, приводящие к изменению инженерно-геологических условий строительства и эксплуатации АЭС и влияющие в основном на работу строящихся объектов.

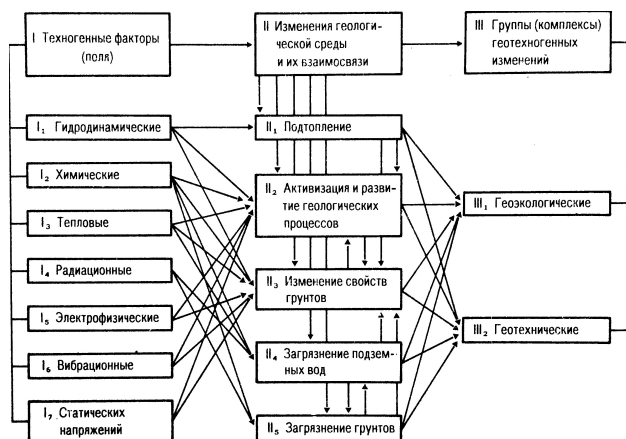


Рис. 3: Влияние техногенных факторов на геологическую среду при строительстве атомных станций

Принципиальное их различие заключается в том, что первые требуют осуществления специальных защитных мероприятий по охране геологической среды, а вторые разработки специальных инженерных решений, обеспечивающих надежность эксплуатации АЭС.

Изменения гидрогеологических условий на площадках АЭС проявляются в изменении положения уровня, химического состава, газовой составляющей, степени минерализации и температуры грунтовых вод. Вследствие этого происходит подтопление территорий, повышаются агрессивность подземных вод и коррозионная способность грунтов, ухудшаются деформационные и прочностные свойства последних,

развивают неблагоприятные геологические процессы (набухание грунтов, просадки суффозия) Изменение положения уровня грунтовых вод (УГВ) может быть вызвано такими причинами, как создание водохранилищ, водопонижение, проходка строительных работ систем котлованов, технического водоснабжения, утечки производственных вод из коммуникаций, нарушение подземного стока в результате засыпки оврагов и других естественных дренажей, барражирование потока грунтовых вод свайными основаниями или фундаментами глубокого заложения, нарушение поверхностного стока, снижение испаряемости вследствие вырубки лесов и застройки территории. Интенсивность повышения УГВ зависит, с одной стороны, от правильности принятых проектных решений и качества строительства, а с другой, от геоморфологического положения площадки, ее геологического строения, гидрогеологических условий и свойств грунтов. Скорость поднятия уровня грунтовых вод составляет 1-2 м/год и может распространяться на площадь радиусом 3-5 км. Наряду с образованием общего купола грунтовых вод на всей строительной площадке формируются локальные повышения (микрокупола) в местах наиболее интенсивных утечек производственных или скопления поверхностных вод. Масштабы подтопления наиболее значительны на площадках водораздельного типа [3].

Кроме повышения УГВ нередко наблюдается формирование верховодки или постоянного водоносного горизонта в приповерхностных отложениях. Этому способствуют факторы, как значительная мощность зоны аэрации, наличие отрицательных техногенных форм рельефа (котлованы, каналы, траншеи), присутствие в разрезе зоны аэрации прослоев и линз грунтов с низким и фильтрационными характеристиками, фильтрационная анизотропия грунтов и т. д.

В отличие от ТЭС на территории атомной станции не происходят накопление золоотвалов и загрязнение атмосферы продукта-

ми газовых выбросов из дымовых труб. В связи с этим изменение минерализации и химического состава подземных вод связано главным образом с выщелачиванием солей из грунтов зоны аэрации в процессе повышения УГВ. Эти процессы нередко весьма интенсивны, особенно в южных районах Восточно-Европейской платформы, где грунты часто сильно засолены. Наибольшую опасность представляет выщелачивание сульфатов, которое резко повышает агрессивность грунтовых вод по отношению к бетону. Однако главная опасность для атомных станций состоит в том, что они могут стать источником радиоактивного загрязнения окружающей среды при возникновении неполадок в работе станции аварийных ситуациях или катастрофах, подобных Чернобыльской источником загрязнения могут оказаться также могильники твердых и жидких отходов, образующихся при работе атомных станций. При этом нужно учитывать, что атомные станции сохраняют экологическую опасность для окружающей среды не только во время их эксплуатации, но и в течение длительного периода времени при ликвидации. Радиоактивные отходы, образующиеся в процессе эксплуатации атомных станций, необходимо надежно изолировать от окружающей среды на длительный срок (по современным представлениям, на 300-500 лет и более) с локализацией их на ограниченном участке литосферы. Саму же станцию после окончания эксплуатации необходимо ликвидировать [2].

Заключение

Радиоактивные вещества, находящиеся в коммуникациях и отходах атомных станций, попадая в подземные воды, мигрируют с ними к местам разгрузки. Таким образом возникает опасность загрязнения подземных вод и дренирующих их поверхностных водоемов, используемых для водоснабжения. В связи с этим при выборе площадок АЭС этот вопрос должен рассматриваться как один из важнейших.

При нормальной работе АЭС утечки радиоактивных компонентов незначительны и не превышают допустимых значений установленных отечественным и международным и нормами. Значительно более сложная ситуация создается при возникновении аварийных ситуаций, когда происходит большой выброс радиоактивных продуктов из реактора и его коммуникаций. В подобных случаях подземные воды и вмещающие их породы следует рассматривать в первую очередь с позиций оценки возможности миграции радиоактивных компонентов и защитной способности геологической среды.

Строительство АЭС приводит к оживлению некоторых геологических процессов и прежде всего карста и карстово-суффозионных явлений. Причиной активизации могут стать утечки производственных вод из различных коммуникации, водопонижение при проходке котлованов или работа водозаборов, инфильтрация воды из водохранилищ, каналов и т. д.

Весьма распространены также просадочные явления в лессовых грунтах, происходящие в результате подъема УГВ. Основных сооружений они не затрагивают, так как фундаменты последних опираются на грунты, залегающие ниже подошвы лессов, но могут повлиять на эксплуатацию вспомогательных сооружений и требуют осуществления защитных мероприятий. Примером может служить Ростовская АЭС.

Как и при гидроэнергетическом строительстве, где закладываются глубокие котлованы, большое внимание должно уделяться изучению выветривания пород, поскольку котлованы остаются открытыми в течение длительного времени, в том числе и зимой разуплотнение и пучение грунтов при промерзании резко снижает их несущую способность, что особенно опасно для терригенных отложений типа аргиллитов и глинистых сланцев. Разуплотнение и снижение прочности морены в открытом котловане было отмечено при строительстве Калининской АЭС. Большую опасность представля-

ют вибрационные воздействия на грунты, особенно после их обводнения.

Литература

- [1] Гаврилов В.П. Общая и историческая геология и геология СССР. – М.: Недра, 1989. — 495 с.
- [2] Сергеев Е.М. Инженерная геология СССР. Платформенные регионы Европейской части СССР. Кн.2. / И.С. Комарова, Д.Г. Зилинга, В.Т. Трофимова. – М.: Недра, 1991. – 357 с.
- [3] Милановский Е.Е. Геология СССР. – М.: Издательство МГУ, 1991. — 272 с.

Исследование влияния воздействия радиоактивных веществ на результаты оценки экологического ущерба вокруг АЭС на примере Калининской АЭС*

А. Ю. Киданова, М. А. Берберова

annakidanova@hotmail.com, maria.berberova@gmail.com

¹Московский физико-технический институт (государственный университет);

²Москва, Международный центр по ядерной безопасности

Несмотря на последствия радиационных аварий последних десятилетий (в первую очередь, аварий на Чернобыльской АЭС в 1986 г. и АЭС Фукусима Дайичи в 2011 г.), а также отказ ряда стран от программ развития ядерной энергетики, оказывается, что на сегодняшний день нет адекватной замены этого направления энергетики на альтернативные: в настоящее время в 30 странах мира насчитывается 195 действующих атомных электростанций (АЭС) с 438 энергоблоками, 71 энергоблок находится в стадии сооружения. Одним из наиболее важных направлений оценки риска является анализ экологического ущерба окружающей среды вокруг АЭС. В целях повышения безопасности управления и эксплуатации АЭС данный анализ необходимо проводить.

Ключевые слова: атомная электростанция, радиоактивное вещество, мощность выброса, радиационно-опасный объект, ущерб.

Investigation of the impact of radioactive substances on the results of the assessment of environmental damage around nuclear power plants using the example of Kalinin NPP*

A. J. Kidanova¹, M. A. Berberova²

¹Moscow Institute of Physics and Technology;

²International nuclear safety center

Despite the effects of radiation accidents in recent decades (primarily the Chernobyl accident in 1986 and the Fukushima Daiichi nuclear power plant in 2011), as well as the refusal of a number of countries to develop nuclear power programs, it turns out that today there is no adequate replacement. This direction of energy for alternative: currently in 30 countries there are 195 operating nuclear power plants (NPPs) with 438 power units, 71 power units are under construction. One of the most important areas of risk assessment is the analysis of environmental damage to the environment around the plant. In order to improve the safety of management and operation of nuclear power plants, this analysis must be carried out.

Key words: nuclear power plant, radioactive substance, emission capacity, radiation-hazardous object, damage.

Введение

В процессе жизнедеятельности взаимодействие человека со средой обитания и её составляющих основывается на передаче между элементами системы потоков веществ и их соединений, энергий всех видов и информации. В техносфере возникают потоки отходов (выбросы в атмосферу, сбросы в водоемы, жидкие и твердые отходы, разнообразные энергетические загрязнения), что в свою очередь может привести к разрушению биохимической структуры среды обитания человека. В связи с этим возникает необходимость определения вероятности проявления негативного воздей-

ствия загрязняющих факторов и размеров их загрязнения [1].

Предмет исследования

Атомная электростанция (на примере Калининской АЭС)

Степень изученности проблемы

На данный момент проведены аналогичные работы для Ростовской АЭС

Цель исследования

Расчет экологического ущерба, вызванного влиянием радиоактивных веществ, вокруг АЭС.

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 16-07-00089, 18-07-20030, 18-37-10008

Задачи исследования

В рамках данного исследования поставлены следующие задачи:

- Разбор существующих методик и их сравнение
- Оценка радиационной обстановки на примере Калининской АЭС
- Оценка экологического ущерба на примере Калининской АЭС

Методологическая основа

На текущий момент известны две методики расчета экологической обстановки вокруг АЭС:

- методика по расчету радиационной обстановки в окружающей среде и ожидаемого облучения населения при кратковременных выбросах радиоактивных веществ в атмосферу (МПА-98)
- методика по разработке и установлению нормативов предельно-допустимых выбросов радиоактивных веществ атомных станций в атмосферный воздух (ПДВ РВ)

Методика МПА-98 содержит рекомендации по расчету радиационной обстановки на местности и ожидаемых доз облучения населения в результате разовых выбросов в атмосферу в произвольные, заранее не запланированные моменты времени. Предполагается, что выброс может произойти в любой, самый неблагоприятный для формирования доз облучения момент времени, поэтому предложенная методология позволяет получить максимально возможные оценки. Данная методика принимает во внимание незапланированные выбросы, при которых не происходит масштабного тепловыделения. При этом облако поднимается на большую высоту. Методика МПА-98 решает следующие две задачи: во-первых, она позволяет рассчитать экстремальные значения факторов разбавления в атмосферу для разовых выбросов и экстремальные значения интегралов концентрации радионуклидов выброса в атмосферу на различных расстояниях от источ-

ника с учетом розы ветров.

Методика ПДВ РВ является улучшением методики МПА-98. Она устанавливает методы расчета предельно допустимых выбросов радиоактивных веществ из специально организованных источников в атмосферу.

В работе рассмотрена каждая из методик отдельно и приведено их сравнение.

Научная новизна

В рамках данной работы поставлена никем ранее не выполненная задача – провести оценку экологической обстановки вокруг Калининской АЭС в результате воздействия радиоактивных веществ и провести оценку экологического ущерба вокруг Калининской АЭС.

Практическая и научная значимость

Любая авария на АЭС множит поводы для формирования негативного общественного мнения по атомной энергетике. Решение задач глобальной ядерной безопасности осуществляется в условиях значительной неопределенности.

Результаты, полученные в данной работе могут быть использованы для корректировки работы АЭС и для получения представления о возможных рисках и их опасности для природы и населения

Границы исследования

В данном исследовании рассмотрен экологический ущерб, связанный с загрязнением почвы и водоемов. Из-за географического расположения Калининской АЭС можно пренебречь экологическим ущербом, связанным с загрязнением леса.

Структура работы

Исследование, проведенное в рамках этой статьи, разделено на три смысловые части.

Первая часть состоит из введения, обозначения цели работы, актуальности работы, ее новизны. Также в этой части присутствует общая характеристика Калининской АЭС, описание состояния территории

расположения Калининской АЭС и воздействие АЭС на окружающую среду.

Во второй части приведен обзор актуальной литературы, рассмотрены методики по расчету экологической обстановки вокруг АЭС, приведено описание каждой из методик, математические формулы, сравнение методик, выведены плюсы и минусы каждой из рассмотренных методик.

В третьей части приведены конкретные расчёты на примере Калининской АЭС, сделаны выводы и заключения.

Общая характеристика Калининской АЭС

Калининская АЭС находится на севере Тверской области в 330 км от Москвы и в 400 км от Санкт-Петербурга. Эта АЭС производит больше всего электроэнергии в центральной части России, здесь вырабатывается почти 75% электроэнергии, потребляемой в Тверской области. Калининская АЭС состоит из четырех энергоблоков, мощность каждого из которых 1000 МВт.

Калининская АЭС включает в себя две очереди: первую (энергоблоки №№ 1 и 2) и вторую (энергоблоки №№ 3 и 4). В состав оборудования каждого из энергоблоков входят:

- водо-водяной энергетический реактор типа ВВЭР-1000;
- парогенератор горизонтального типа ПГВ-1000;
- паровые турбины типа К-1000-60/1500 ПОАТ «ХТЗ» (на блоках 1 и 2) и турбины К-1000-60/3000 ПО «ЛМЗ» (на блоках 3 и 4);
- генераторы типа ТВВ-1000-4УЗ (1-я очередь) и генераторы типа ТВВ-1000-2УЗ (2-я очередь);
- главные циркуляционные насосы типа ГЦН-195М.

На рис. 1 представлена технологическая схема работы энергоблока с реактором ВВЭР-1000:

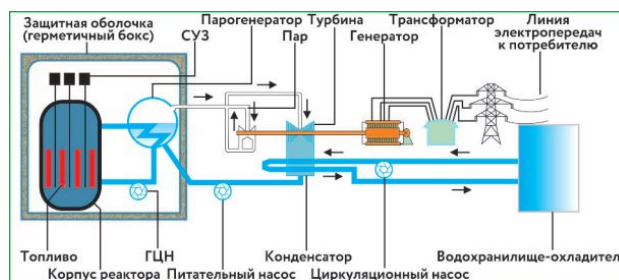


Рис. 1: Технологическая схема работы энергоблока с реактором ВВЭР-1000.

Согласно [5], в соответствии с международными стандартами в области охраны окружающей среды [6, 7] на Калининской АЭС в 2009 году была принята «Экологическая политика» ОАО «Концерн Росэнергоатом». Документ введен в действие приказом Калининской АЭС от 21.09.2009 № 1112. В связи с организационными изменениями в концерне в 2012 году, приказом от 21.09.2012 № 107-П/вн была произведена корректировка выпущенного ранее приказа. В связи с пересмотром в 2014 году «Экологической политики» ОАО «Концерн Росэнергоатом» на Калининской АЭС приказом от 19.02.2014 № 381-П/вн была введена актуализированная «Экологическая политика» ОАО «Концерн Росэнергоатом» [8].

Целью «Экологической политики» [8] является обеспечение устойчивого экологически ориентированного развития атомной энергии и поддержания такого уровня безопасности АС, при котором воздействие на окружающую среду, персонал и население на ближайшую перспективу и в долгосрочном периоде обеспечивает сохранение природных экосистем, поддержание их целостности и жизнеобеспечивающих функций.

Основные принципы «Экологической политики» [8]:

- принцип соответствия;
- принцип предупреждения;
- принцип готовности;
- принцип информационной открытости;
- принцип системности.

Принцип соответствия заключается в соблюдении соответствия производствен-

ной деятельности законодательным и другим нормативным требованиям и стандартам, в том числе международным, в области обеспечения безопасности и охраны окружающей среды, неукоснительное выполнение каждым работником норм и правил в области обеспечения безопасности персонала и населения и охраны окружающей среды.

Принцип предупреждения негативного воздействия - система приоритетных действий, направленных на недопущение опасных экологических аспектов, которые могут оказать негативное воздействие на человека и окружающую среду.

Принцип готовности заключается в постоянной готовности руководства и персонала к предотвращению техногенных аварий и иных чрезвычайных ситуаций и ликвидации их последствий.

Принцип информационной открытости заключается в прозрачности и доступности экологической информации.

И последний принцип, **принцип системности**, подразумевает системное и комплексное решение вопросов обеспечения экологической безопасности, целевого планирования и ведения природоохранной деятельности с учетом многофакторности аспектов безопасности на основе современных концепций анализа рисков и экологических ущербов.

Состояние территории расположения Калининской АЭС

В 30-километровую зону вокруг АЭС входят 49 охраняемых территорий, из них 16 памятников природы и 33 заказника. Это является свидетельством сохраняемого биоразнообразия и стабильности экосистем, минимального влияния негативных производственных факторов предприятия на окружающую среду. На территории Удомельского района зарегистрировано более 220 видов птиц, отмечено 911 видов растений, из которых 68 включены в Красную книгу Тверской области, 7 видов занесены в Красную книгу РФ. Постоянно иду-

щие мониторинговые исследования наземных и водных экосистем регулярно пополняют список зарегистрированных в окрестностях КЛН АЭС редких растений, насекомых, птиц. Исследования состояния наземных экосистем, в том числе почвенного покрова, растительности, животного мира проводили сотрудники кафедры биологии Тверского государственного университета. В целом на пунктах мониторинга была отмечена высокая стабильность характеристик почвенного покрова, типичная естественная динамика наблюдаемых экосистем, что является свидетельством отсутствия значимого влияния производственных факторов Калининской АЭС на их функционирование. Причиной же наблюдаемых некоторых изменений является общее повышение антропогенной нагрузки.

Воздействие Калининской АЭС на окружающую среду

В комплексную программу экологического мониторинга района и площадки Калининской АЭС входят:

- гидрологические наблюдения;
- метеорологические наблюдения;
- наблюдения за уровнем подземных вод на промплощадке АЭС 1 и 2 очереди;
- наблюдения за осадками фундаментов и деформациями зданий и сооружений 2 очереди;
- наблюдения за микродеформациями грунтового основания энергоблоков 3 и 4;
- наблюдения за современными движениями земной коры и гравитационным полем района и площадки;
- сейсмометрические наблюдения на геодинамическом полигоне;
- программа экологического мониторинга наземных и водных экосистем региона Калининской АЭС;
- мониторинг здоровья населения.

Санитарно-защитная зона Калининской АЭС установлена распоряжением администрации Удомельского района радиусом в 1.2 км, отсчитываемым от геометрического центра вентиляционных труб всех четырех

энергоблоков. Дополнительно в нее включена территория под сбросной канал и градирни. Зона наблюдения Калининской АЭС составляет окружность радиусом 11 км. Другим важнейшим видом контроля является контроль мощности дозы гамма-излучения на местности, который осуществляется 18 мониторинговыми станциями автоматизированной системы контроля радиационной обстановки (АСКРО), установленными в 30-ти километровой зоне расположения Калининской АЭС. В АСКРО Калининской АЭС используются многоканальные измерительные установки. Данные о состоянии радиационной обстановки, метеоданные и др. автоматически передаются от гамма-зондов радиосигналом в центральные посты, размещенные на станциях. По результатам контроля мощности дозы гамма-излучений по всем установкам АСКРО Калининской АЭС в 2015 году изменений уровней естественного фона в сторону увеличения не зарегистрировано.

Отдельно стоит рассмотреть **сбросы радионуклидов**. Индекс сброса радионуклидов (сумма отношений активности радионуклида к допустимому сбросу) с жидкими стоками Калининской АЭС в 2015 году составил: 0.019, что гарантирует непревышение установленной СП АС-03 [12] квоты на облучение населения 10 мкЗВ/год.

Год	2014	2015
Объем сброса в м ³	73872	62198
Индекс сброса	0.015	0.019

Также стоит обратить внимание на Выбросы вредных химических веществ. На основной производственной площадке Калининской АЭС имеются две установки очистки газов. Эффективность очистки на основании выполненных измерений составляет 93,9%. Например, на 77 источников выброса в 2015 году пришелся 21 неорганизованный источник выброса.

Обзор методик по расчету экологической обстановки вокруг АЭС

На текущий момент известны две методики расчета экологической обстановки вокруг АЭС:

- методика расчета радиационной обстановки в окружающей среде и ожидаемого облучения населения при кратковременных выбросах радиоактивных веществ в атмосферу (методика МПА-98) [2-4, 10];
- методика разработки и установления нормативов предельно-допустимых выбросов радиоактивных веществ атомных станций в атмосферный воздух (методика ПДВ РВ) [2-4, 11].

Методика МПА-98 [2-4, 10] указывает, как рассчитывать радиационную обстановку в окружающей среде и облучение населения при кратковременных выбросах радиоактивных веществ в атмосферу. Данная методика принимает во внимание незапланированные выбросы, при которых не происходит масштабного тепловыделения. При этом облако поднимается на большую высоту. Методика МПА-98 решает следующие две задачи: во-первых, она позволяет подсчитывать экстремальные значения факторов разбавления в атмосферу для разовых выбросов и экстремальные значения интегралов концентрации радионуклидов выброса в атмосферу на различных расстояниях от источника с учетом ветра.

Методика ПДВ РВ [2-4, 11] является улучшением методики МПА-98 [2-4, 10]. Она устанавливает методы расчета предельно допустимых выбросов радиоактивных веществ из специально организованных источников в атмосферу.

Рассмотрим каждую из методик отдельно.

Методика МПА-98

В методике МПА-98 [2-4, 10] используется гауссовская модель атмосферной диффузии. Пусть:

- $C_{v,\rho}^o$ - концентрация выбрасываемого радионуклида ρ в атмосфере

- $Q_{o,\rho}$ - величина мгновенного выброса радионуклида ρ
- $G'_{o,\rho}(x, t)$ - фактор разбавление мгновенного выброса в атмосфере в m^{-3}
- $h(x)$ - высота центра облака над поверхностью земли в метрах
- $F(x)$ - фактор истощения облака выброса на расстоянии x
- $\sigma_x, \sigma_y, \sigma_z$ - дисперсии облака выброса по направлениям движения облака, где x - движение по ветру, y и z - движения поперек ветра

Тогда эта концентрация следующим образом зависит от расстояния x от источника выброса:

$$C_{v,\rho}^o = Q_{o,\rho} * G'_{o,\rho}(x, t)$$

Причем

$$G'_{o,\rho}(x, t) = \frac{2}{(2\pi)^{3/2}} \frac{F_\rho(x)}{\sigma_x \sigma_y \sigma_z} e^{-\frac{(x-ut)^2}{2\sigma_x^2}} e^{-\frac{h(x)^2}{2\sigma_z^2}}$$

Тогда временной интеграл фактора разбавления равен

$$G_{o,\rho}(x) = \int_t G'_{o,\rho}(x, t) = \frac{F_\rho(x)}{\pi \sigma_y \sigma_z u} e^{-\frac{h(x)^2}{2\sigma_z^2}}$$

Для выполнения расчетов нужно определить значения $\sigma_x, \sigma_y, \sigma_z$. Поперечные σ_y, σ_z рассчитываются по формуле Эйри:

$$\sigma(x, z_0, P) = \frac{x}{\sqrt{a(z_0, P) + x * b(z_0, P)}}$$

где $a(z_0, P)$ и $b(z_0, P)$ - коэффициенты, которые зависят от P (параметра устойчивости атмосферы), и z_0 (высоты шероховатости поверхности). Значения P и z_0 представлены в таблицах 1 и 2.

Таблица 1: Значения параметра P

Категория устойчивости	Значения параметра P
A	0.5
B	1.5
C	2.5
D	3.5
E	4.5
F	5.5
G	6.5

Таблица 2: Значения переменной z_0

Микрорельеф	Значения z_0 в см
Снег	0.1
Газон до 1см	0.1
Трава до 15см	0.6-2
Высокая трава до 60см	4-9
Пов-ть с участками травы	10-20
Парк, лес	20-100
Город	400

Пусть:

- λ_ρ - постоянная радиоактивного распада радионуклида ρ в c^{-1}
- u - скорость ветра в центре облака в м/с
- Λ - постоянная вымывания осадками в c^{-1}

Рассмотрим две функции истощения облака. Функция истощения облака выброса в результате радиоактивного распада равна

$$F_\rho(x) = e^{-\lambda_\rho \frac{x}{u}} = e^{-\sqrt{\frac{2}{\pi}} \frac{V_g}{u} \int_0^x \frac{1}{\sigma_z(x)} e^{-\frac{h^2}{2\sigma_z^2}} dx}$$

Значения параметра V_g можно найти в таблице 3.

Таблица 3: Значения параметра V_g

Вещество	Значения V_g в м/с
Аэрозоли	0.8
Элементарный йод	2
Органические соединения йода	0.1

Функция истощения облака выброса в результате влажного выведения равна

$$F_\Lambda(x) = e^{-\Lambda \frac{x}{u}}$$

При этом

$$\Lambda = k_\rho k_o I,$$

где

I - интенсивность осадков в мм/ч,

k_ρ - величина абсолютной вымывающей способности дождя в $ч/(мм \cdot с)$,

k_o - относительная вымывающая способность осадков (значения переменной k_o представлены в таблице 4.

Таблица 4: Значения переменной k_o

Тип осадков	Значения k_o
Дождь	1.0
Дождь с грозой	1.1
Снег с дождем	2.4
Ливень	2.8
Снег	3.0
Морозь	4.5
Туман	5.0

Методика ПДВ РВ

Данная методика [11] является улучшением технической методики [10].

Методика устанавливает методы проведения расчета нормативов предельно допустимых выбросов радиоактивных веществ из организованных источников в атмосферный воздух.

Пусть P - норма предельно допустимых выбросов.

Ниже представлена формула среднегодового метеорологического фактора разбавления радионуклида ρ в приземном слое:

$$G_{\rho,n}(x) = \frac{2N}{2\pi^{3/2}R_n x} \sum_j \sum_k \frac{\omega_{n,j,k} F_{\rho,j,k}}{\sigma_{z,j}(x) U_{j,k}} e^{-\frac{h(x)^2}{2\sigma_{z,j}(x)^2}},$$

где:

- 1) деление на два в знаменателе степени экспоненты берется согласно формуле нормального распределения;
- 2) формула учитывает, что вклад дисперсии облака поперек ветра незначителен по сравнению с вкладом дисперсии по оси z ;
- 3) $\sum_n \sum_k \sum_j \omega_{n,j,k} = 1$, $\sigma_{z,j}(x)^2$ - дисперсия струи по вертикали;
- 4) j - категория устойчивости;
- 5) $\omega_{n,j,k}$ - повторяемость метеорологических условий;
- 6) R_n - поправка на рельеф;
- 7) $F_{\rho,j,k}$ - фактор истощения струи;
- 8) $h(x)$ - геометрическая высота трубы в метрах;
- 9) n - номер румба;
- 10) $U_{j,k}(x) = U_k(z_\varphi) \frac{h}{z_\varphi} b_j$ - модуль скорости ветра в м/с;
- 11) $U_k(z_\varphi)$ - модуль приземной скорости ветра (k - градация по скоростям);
- 12) z_φ - высота флюгера;

13) b_j - параметр, выражающийся следующим образом $b_j = a_1 + a_2 z_0^{a_3}$.

Значения коэффициентов a_1, a_2 представлены в таблице 5.

Таблица 5: Значения параметров a_1, a_2

Категория устойчивости	a_1	a_2
A	0.112	5.38e-4
B	0.13	6.52e-4
C	0.112	9.05e-4
D	0.098	1.35e-3
E	0.08	1.58e-3
F	0.061	1.96e-3
G	0.064	1.36e-3

Дисперсионные кривые можно представить в виде:

$$\sigma_z(x) = f(z_0, x)g(x)$$

при

$$f(z_0, x)g(x) < \sigma_{max}$$

и

$$\sigma_z(x) = \sigma_{max}$$

при

$$f(z_0, x)g(x) \geq \sigma_{max}, \text{ где } f(z_0, x) = \ln(c_1 x^{d_1} (1 + c_2 x^{d_2})^{-1}) \text{ и } g(x) = \frac{a_1 x^{b_1}}{1 + a_2 x^{b_2}},$$

где

a_1, a_2, b_1, b_2 - параметры категории устойчивости атмосферы, значения этих коэффициентов представлены в таблицах 5 и 6, c_1, c_2, d_1, d_2 - коэффициенты шероховатости, значения этих коэффициентов представлены в таблицах 7 и 8.

Значения параметров a_1, a_2, b_1, b_2 зависят от категории устойчивости атмосферы, а значения c_1, c_2, d_1, d_2 - от коэффициента шероховатости.

Таблица 6: Значения параметров b_1, b_2

Категория устойчивости	b_1	b_2
A	1.06	0.815
B	0.95	0.755
C	0.92	0.718
D	0.889	0.688
E	0.892	0.686
F	0.895	0.684
G	0.783	0.672

Таблица 7: Значения параметров c_1, c_2

z_0	c_1	c_2
0.01	1.56	0.048
0.04	2.02	0.027
0.1	2.72	0
0.4	5.16	-0.098
1.0	7.37	-0.096

Таблица 8: Значения параметров d_1, d_2

z_0	d_1	d_2
0.01	6.25e-4	0.45
0.04	7.76e-4	0.37
0.1	0	0
0.4	18.6	-0.225
1.0	4.29e-3	-0.6

$\omega_{n,j,k}$ можно найти следующим образом:

$$\omega_{n,j,1} = \frac{m_{n,j,1}}{M} = \frac{\sum_{j=1}^J m_{n,j,2}}{\sum_{n=1}^N \sum_{j=1}^J m_{n,j,2}},$$

$$\omega_{n,j,k>1} = \frac{m_{n,j,k>1}}{M},$$

где

$M = \sum_{n=1}^N \sum_{j=1}^J \sum_{k=1}^K m_{n,j,k}$ - общее число наблюдений за состоянием атмосферы.

Выбросы радионуклидов в атмосферных воздушных потоках не приведут к ограничению использования поверхностного слоя почвы, если в любой точке местности выполняется условие

$$\sum_{\rho} \frac{P^{\rho}(D_{g,n^{\rho}} + D_{w,n^{\rho}})}{U^{\rho} \lambda_{eff}^{\rho} r h} \leq 1,$$

где

U^{ρ} - удельная активность радионуклида ρ ,

r - плотность поверхностного слоя почвы,

h - толщина поверхностного слоя почвы,

$D_{g,\rho} = V_{g,\rho} G_{\rho,n}$ - среднегодовой метеорологический фактор сухого осаждения,

$D_{w,n} = \lambda \int_0^{\text{inf}} G_{r,n}(x, z) dz$ - среднегодовой метеорологический фактор влажного выведения.

Далее мы можем определить критическое значение P , при котором можно в дальнейшем использовать природные ресурсы.

Сравнение предложенных методик

Важным преимуществом методики ПДВ РВ [11] перед методикой МПА-98 [10] явля-

ется тот факт, что методика ПДВ РВ учитывает больше параметров, следовательно, дает более точные результаты. Методика МПА-98 [10], как и методика ПДВ РВ [11], обладает следующими достоинствами: она учитывает устойчивость атмосферы, шероховатость поверхности и все виды истощения облака. Большим недостатком методики МПА-98 [10] является ее неприменимость при долговременном наблюдении за состоянием атмосферы и при наблюдениях на больших расстояниях

Расчеты для Калининской АЭС:

В данном исследовании рассмотрен экологический ущерб, связанный с загрязнением почвы и водоемов. Из-за географического расположения Калининской АЭС можно пренебречь экологическим ущербом, связанным с загрязнением леса. Поэтому для Калининской АЭС экологический ущерб рассчитывается по следующей формуле:

$$M_g + M_w$$

где M_g - выплаты за ущерб, связанный с загрязнением почвы, M_w - выплаты за ущерб, связанный с загрязнением водоемов

$$M_g = S_g h p C_g$$

где

S_g - площадь плодородного слоя почвы,

h - глубина утраченного слоя почвы,

p - средняя плотность почвы.

C_g - рыночная стоимость тонны земли

$$M_w = 5 \sum_i S_i W_i K_e K_p K_{inf}$$

где

S_i - выплаты за сброс тонны загрязняющего вещества в водоем,

W_i - масса выброшенного вещества,

K_e - коэффициент экологической ситуации водоема,

K_p - коэффициент особо охраняемых территорий,

K_{inf} - инфляция.

Для задачи расчета радиационной обстановки за год рассматривались более распространенные радионуклиды при выбросе:

- Co^{60} - 0.1%
- Cs^{134} - 0.2%
- Cs^{137} - 0.2%
- I^{131} - 5%
- ИРГ - 4%

С наибольшей частотой происходит выброс I^{131} .

Сравним полученный результат с результатом расчета [2] для Ростовской АЭС. В [2] получено, что:

- Sr^{90} - 20%
- Co^{60} - 24%
- Cs^{137} - 27%
- I^{131} - 28%

Здесь также с наибольшей частотой происходит выброс I^{131} .

На рис. 2 для I^{131} представлен график зависимости среднегодового метеорологического фактора разбавления от расстояния до источника выброса при восьми различных румбах ветра [13]:

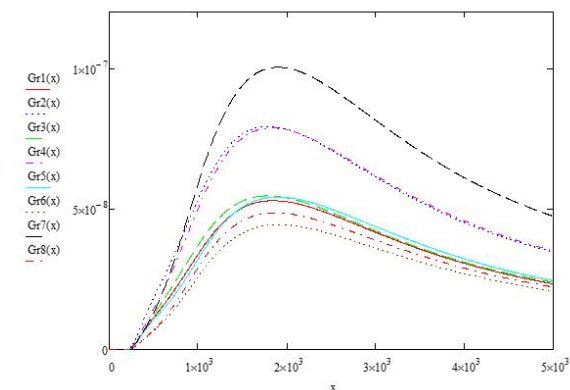


Рис. 2: Технологическая схема работы энергоблока с реактором ВВЭР-100: $Gr1(x)$ - север, $Gr2(x)$ - северо-восток, $Gr3(x)$ - восток, $Gr4(x)$ - юго-восток, $Gr5(x)$ - юг, $Gr6(x)$ - юго-запад, $Gr7(x)$ - запад, $Gr8(x)$ - северо-запад

Для Co^{60} , Cs^{134} , Cs^{137} и для I^{131} в том числе максимального среднегодового метеорологического фактора разбавления от расстояния до источника выброса при восьми раз-

личных румбах ветра представлены в таблицах 9 и 10.

Таблица 9: Значения максимальных значений среднегодового фактора разбавления Co^{60} и Cs^{134}

направление	Co^{60}	Cs^{134}
1 - С	5,784	5.096
2 - СВ	3,836	3.006
3 - В	8,597	7.992
4 - ЮВ	6,051	5.93
5 - Ю	5,29	5.048
6 - ЮЗ	7,912	6.027
7 - З	5,382	4.012
8 - СЗ	7,781	7.088

Таблица 10: Значения максимальных значений среднегодового фактора разбавления Cs^{137} и I^{131}

направление	Cs^{137}	I^{131}
1 - С	5,409	5,155
2 - СВ	3,361	4,231
3 - В	9,954	9,577
4 - ЮВ	4,651	4,654
5 - Ю	5,289	5,044
6 - ЮЗ	7,95	7,536
7 - З	5,502	5,2
8 - СЗ	7,87	7,524

Из таблиц 9 и 10 видно, что максимальная концентрация радионуклидов выпадает на восток, это объясняется тем, что на Калининской АЭС преобладают западные и юго-западные ветра.

Заключение

В работе проведено исследование двух методик расчета экологической обстановки вокруг АЭС: методики по расчету радиационной обстановки в окружающей среде и ожидаемого облучения населения при кратковременных выбросах радиоактивных веществ в атмосферу (МПА-98) и методики по разработке и установлению нормативов предельно-допустимых выбросов радиоактивных веществ атомных станций в атмосферный воздух (ПДВ РВ). Рассмотрены достоинства и недостатки этих методик.

В работе выполнен расчет экологического ущерба для Калининской АЭС, связанного с загрязнением почвы и водоемов. Проведено сравнение полученных результатов с результатами выполненного ранее расчета для Ростовской АЭС.

Выводы

Проведенные в данной работе и в [2] расчеты экологического ущерба, соответственно для Калининской и Ростовской АЭС показали, что с наибольшей частотой происходит выброс I^{131} . Существенные различия в результатах расчетов можно объяснить разным типом местности: Калининская АЭС - это леса, болота и пр., Ростовская АЭС - это, прежде всего, степь.

Так же, как и в [2], не решена проблема оценки радиационной обстановки в плоскости поверхности не только по направлениям розы ветров, но и во всех других направлениях. Данная оценка даст более точные результаты радиационной обстановки, что, в свою очередь, улучшит точность оценки экологического ущерба.

И так же, как и в [2] не изучены способы расчета ущерба атмосфере, лесопарковой зоне и водной структуре.

Литература

[1] Марунин, В.И., Грузин, Д.П. Анализ экологического риска / В.И.Марунин, Д.П.Грузин // Надежность и качество. Т. 2: труды международного симпозиума. - Пенза: Изд. ПГУ, 2010. - С. 278.

[2] Илясов П.С. Оценка экологического ущерба вокруг АЭС в результате воздействия радиоактивных веществ (на примере Ростовской АЭС): диссертация на соискание ученой степени магистра: 010900 / Илясов Петр Сергеевич. - М., 2017. - 36 с

[3] Исламов Р.Т., Берберова М.А., Илясов П.С. Разработка методического подхода к оценке экологического ущерба вокруг АЭС в результате воздействия радиоактивных веществ / М.А. Берберова, П.С.

Илясов, Р.Т. Исламов // SCVRT1516: труды Международной научной конференции. - Протвино-Москва: Изд. ИФТИ, 2016, С. 214-219.

[4] Исламов Р.Т., Берберова М.А., Илясов П.С. Оценка экологического ущерба вокруг АЭС в результате воздействия радиоактивных веществ / Р.Т. Исламов, М.А. Берберова, П.С. Илясов // СРТ1617: труды Международной научной конференции. - Протвино-Москва: Изд. ИФТИ, 2017, С. 197-201.

[5] Отчет об экологической безопасности за 2013 год - М.: Филиал ОАО "Концерн "Росэнергоатом"Калининская АЭС, 2013- 32 с.

[6] ГОСТ Р ИСО 14001—2007 Системы экологического менеджмента. Требования и руководство по применению. - М., 2007. - 27 с.

[7] ГОСТ Р ИСО 14004—2007 Системы экологического менеджмента. Руководящие указания по принципам, системам и методам обеспечения функционирования» Общее руководство по принципам, системам и методам обеспечения функционирования. - М., 2007. - 42 с.

[8] Экологическая политика ОАО «Концерн Росэнергоатом»: [утверждена и введена в действие приказом ОАО «Концерн Росэнергоатом» от 3 февраля 2014 г. № 9/96-П]. - М., 2014. - 2 с.

[9] Отчет по экологической безопасности за 2015 год - М.: Акционерное общество "Российский концерн по производству электрической и тепловой энергии на атомных станциях 2015 - 36 с.

[10] Методические указания по расчету радиационной обстановки в окружающей среде и ожидаемого облучения населения при кратковременных выбросах радиоактивных веществ в атмосферу: [Методические указания МПА-98: утверждены приказом Министра РФ по атомной энергии от 10.01.1999 г. № 12]. - М., 1999. - 126 с.

[11] Методика по разработке и установлению нормативов предельно-допустимых

выбросов радиоактивных веществ атомных станций в атмосферный воздух: [Методика МТ 1.2.1.15.1176-2016: утверждена приказом АО «Концерн Росэнергоатом» от 29.12.2016 г. № 9/1786-П]. - М., 2016. - 70 с.

[12] Санитарные правила проектирования и эксплуатации атомных станций (СП АС-03): [Санитарные правила и гигиенические нормативы СанПин 2.6.1.24-03: утвержде-

на Главным государственным санитарным врачом РФ 22 апреля 2003 г.]. - М., 2003. - 81 с.

[13] Киданова А.Ю., Берберова М.А. Анализ экологического ущерба вокруг Калининской АЭС в результате воздействия радиоактивных веществ / А.Ю.Киданова, М.А.Берберова // SCVRT2017: труды Международной научной конференции. - Протвино-Москва: Изд. ИФТИ, 2017, С. 74-79

Исследование влияния стресса на вероятность совершения ошибок персоналом АЭС в рамках анализа надежности персонала*

А. А. Сынгаевская¹, М. А. Берберова², А. А. Деревянкин²
syngaevskaya@phystech.edu

¹Московский физико-технический институт (государственный университет);

²Москва, Международный центр по ядерной безопасности

Стресс является неотъемлемым компонентом повседневной жизни и профессиональной деятельности персонала АЭС. Вклад человеческого фактора в аварийные ситуации на объектах техносферы существен: 70% авиакатастроф, 50% катастроф на флоте за последние годы произошли из-за неправильных действий (низкой надежности) персонала. Статистические данные свидетельствуют, что основными причинами аварий являются неправильные действия (низкая надежность) персонала (60-70%), технические причины (20-30%), неблагоприятное воздействие внешних факторов и др. [1]. Таким образом, важнейшим условием безаварийной работы потенциально опасных объектов является обеспечение надежности их персонала. На надежность персонала влияет совокупность эмоциональных, волевых, мотивационных, интеллектуальных и других личностных качеств, обеспечивающих точное, безошибочное, адекватное восприятие сложившейся ситуации, своевременное и успешное выполнение регламентированных функций в различных режимах работы. Все эти факторы влияют на способность оператора своевременно и корректно принимать верные цепочки принятия решений при работе с оборудованием. В статье рассматриваются основные определения и понятия анализа надежности персонала (АНП). Вследствие аварий возникает возможность не только отказа оборудования, но и нанесения ущерба персоналу, населению, проживающему вблизи АЭС, и экологической обстановке не только вблизи АЭС, но и на достаточно большом расстоянии от нее.

Ключевые слова: стресс, анализ надежности персонала, вероятностный анализ безопасности, ошибки персонала.

Analysis of the impact of stress on the probability of errors committed by NPP personnel within the framework of staff reliability analysis*

А. А. Synhaevskaya¹, М. А. Berberova², А. А. Derevyankin²

¹Moscow Institute of Physics and Technology (State University);

²Moscow, International nuclear safety center

Stress is an integral component of the daily life and professional activities of NPP personnel. The contribution of the human factor to emergency situations at the facilities of the technosphere is significant: 70% of air crashes, 50% of catastrophes in the fleet in recent years have been due to incorrect actions (low reliability) of personnel. Statistics show that the main causes of accidents are incorrect actions (low reliability) of personnel (60-70%), technical reasons (20-30%), adverse effects of external factors, etc. [1]. Thus, the most important condition for trouble-free operation of potentially hazardous facilities is to ensure the reliability of their personnel. The reliability of personnel is affected by a combination of different personal qualities like: emotions, will, motivation and intellect. These qualities provide accurate, error-free, adequate perception of the situation, timely and successful performance of regulated functions in various modes of operation. All these factors influence personal's ability to make quick and correct decisions while working with equipment. The main definitions and concepts of personnel reliability analysis (ANP) are considered in the article. Due to accident the possibility of equipment failure, personnel and nearby population injury arises. At the same time the environmental situation near the NPP and sometimes on significant distances from NPP is jeopardised.

Keywords: stress, personnel reliability analysis, probabilistic safety analysis, errors

Введение

Надежность человека-оператора является важнейшим показателем безопасной эксплуатации АЭС. Ошибки могут совершать-

ся при проведении проверок работоспособности, технического обслуживания, на этапе управления аварией и т. д. На стабильность работы АЭС влияет ряд различных факторов [1, 2]:

- уровень организации проекта;
- качество эксплуатируемого оборудования;

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 16-07-00089, 18-07-20030, 18-37-10008

- отбор и подготовка компетентного персонала;
- поддержание квалификации работников АЭС.

и др.

В системе «человек-машина» надежность технической составляющей рассчитывается известными методами и в соответствии с установленными стандартами надежности. Однако «человеческая» составляющая не может быть технически точно определена, поэтому необходимо предпринимать системные усилия по повышению и последующему поддержанию достигнутого уровня надежности этой составляющей.

Необходимость учета человеческого фактора при проведении анализа безопасности и надежности АЭС обоснована тем фактом, что ошибочные действия персонала и операторов АЭС прямо или косвенно могут привести к возникновению аварий. Поэтому выполнение АНП при проведении вероятностного анализа безопасности (ВАБ) и оценок риска позволяет выявить наиболее вероятные ошибочные действия персонала АЭС и разработать комплекс мер по их снижению.

В зависимости от вероятных последствий аварии и времени на вмешательство, которым располагает оператор, персонал будет подвергаться различным уровням стресса, что напрямую повлияет на выполнение действий персонала по устранению неполадок при возникновении аварийной ситуации. Если авария произошла, то за короткий промежуток времени оператору необходимо принять ряд решений, корректность которых может как спасти, так и усугубить ситуацию.

Надежность персонала

Условием безаварийной работы потенциально опасных объектов является обеспечение высокой надежности персонала. На надежность персонала влияет совокупность эмоциональных, волевых, мотивационных,

интеллектуальных и других личностных качеств, обеспечивающих точное, безошибочное, адекватное восприятие сложившейся ситуации, своевременное и успешное выполнение регламентированных функций в различных режимах работы. Все эти факторы влияют на способность оператора своевременно и корректно принимать верные решения при работе с оборудованием. Целью работы является исследование влияния стрессовых (аварийных) ситуаций на процесс принятия решения по устранению аварии. Также планируется изучить зависимость стрессоустойчивости от стажа персонала АЭС.

Некоторые заключения исследований, проводимых на тренажере ВВЭР-440 Нововоронежского учебного центра [3]:

1. Операторы ухудшали свои показатели при переходе от нормальных режимов работы к аварийным. При этом у опытных операторов снижение показателей значительно ниже, чем у стажеров.
2. Именно опытные операторы чаще совершали грубые ошибки, приводящие к окончательному невыполнению функций.
3. Более 30% стажеров характеризуются хорошими показателями при нормальной эксплуатации и весьма невысокими при авариях.

Анализ ошибок персонала

Целью АНП является анализ ошибок персонала (ОП). Ошибки могут совершаться при проведении проверок работоспособности, технического обслуживания, на этапе управления аварией и т. д. В АНП необходимо учитывать ряд факторов, таких как: качество инструкций, уровень подготовки персонала и др. [1, 2].

Анализ надёжности персонала — это оценка риска, связанная с действиями персонала. Количественная оценка вероятностей ошибок персонала (ВОП), в основном, имеет качественный характер и трудно поддается

ся статистическому оцениванию на основе эксплуатационных данных.

В вероятностном анализе безопасности (ВАБ) ошибка персонала — это событие, которое представляет собой невыполнение или некорректное выполнение действий (единичных или множественных), в том числе вследствие принятия неправильных решений. Для оценки риска связанного с ошибками персонала важно определить корректный метод оценки вероятности ошибок. Существует большое количество методов, применяемых для АНП, например, таких как метод дерева решений, THERP (Technique for Human Error Rate Prediction), SLIM (Success Likelihood Index Methodology) и др. [1, 2].

Ошибки персонала

Существуют следующие категории ошибок персонала (ОП), моделируемые в ВАБ [1, 2]:

1. **До-аварийные ОП.** Представляют собой действия, совершаемые до возникновения исходного события и приводящие к неработоспособности оборудования или системы, обнаруживаемые только в момент возникновения требования на их функционирование в ответ на исходное событие (ИС).
2. **ОП, иницирующие ИС.** Представляют собой ошибки при выполнении действий персонала, которые приводят к ИС.
3. **Послеаварийные ОП.** Представляют собой невыполнение действий, являющихся ответными на аварийную ситуацию после наступления ИС. Возникают вследствие ошибок диагностики и ошибок выполнения.
4. **ОП при выполнении восстанавливающих действий.** ОП, представляющие невыполнение действий на запуск оборудования вручную (при отказе автоматики) или на организацию альтернативных путей выполнения отказавшей функции.

ОП, моделируемые в ВАБ, обычно подразделяются на следующие виды (механизмы), относящиеся к различным типам поведения [1, 2]:

- оплошность/упущение — действие, приводящее к непреднамеренному результату из-за ошибки при выполнении. Как правило, этот тип ошибок относится к действиям, основанным на навыке;
- заблуждение — намеренно выполненное действие, но намерение является неверным. Это типичная ошибка при оценке ситуации или при планировании ответных действий, приводящая к выполнению неправильных действий. Этот вид ошибок относится к действиям, основанным на правилах или знаниях.

ОП, моделируемые в ВАБ, также подразделяются на ошибки невыполнения и ошибки выполнения:

- ошибка невыполнения — ошибка персонала, заключающаяся в невыполнении одного или более предписанных действий;
- ошибка выполнения — ошибка персонала, заключающаяся в неправильном выполнении одного или более действий, или выполнении какого-либо другого действия/действий вместо предписанных.

Зависимость между действиями персонала имеет место, если результат конкретного действия зависит от предыдущего действия или действий. В соответствии с, вероятности ОП при зависимых действиях должны быть выше, чем, если бы они рассматривались независимыми. Все аварийные последовательности, которые включают несколько ОП, анализируются на предмет их потенциальных зависимостей. По возможности, ОП, которые полностью зависимы, переопределяются и моделируются как единое событие.

Таблица 1: Области вероятностей ошибок с учетом типов поведения

Типы поведения оператора	Области вероятностей ошибок
Действия, основанные на опыте	$P_r = 10^{-4} - 10^{-2}$
Действия, основанные на правилах (письменных процедурах)	$P_r = 10^{-3} - 10^{-1}$
Действия, основанные на знаниях	$P_r = 10^{-2} - 1$

Типы поведения оператора

Выделяются следующие типы поведения оператора при выполнении действий, направленных на устранение последствий аварии [1, 2]:

- Основанные на навыке.** Действия, отработанные до автоматизма по сценариям, изложенным в инструкциях.
- Основанные на правилах.** Действия по правилам менее отработаны, чем действия, основанные на навыке, т. к. выполняются реже и являются более сложными и зачастую требуют повышенного осознанного контроля. Ошибки при этом типе действий связаны с неправильной диагностикой ситуации, приводящей к использованию неверного правила или инструкции.
- Основанные на знаниях.** Действия в незнакомых ситуациях, когда имеющиеся правила и опыт не могут быть применены непосредственно. Для этих действий не существует готовых инструкций, они основаны на общих знаниях, и выполнять которые необходимо в режиме реального времени. Ошибки при этом типе действий являются следствием ограниченности времени на принятие решения и выполнения действия, а также недостаточности или неверности знаний.

Согласно Рассмуссену [2, 4], на практике бывает сложно определить четкую границу между действиями персонала основанных на навыке и на правилах. В таблице 1 пред-

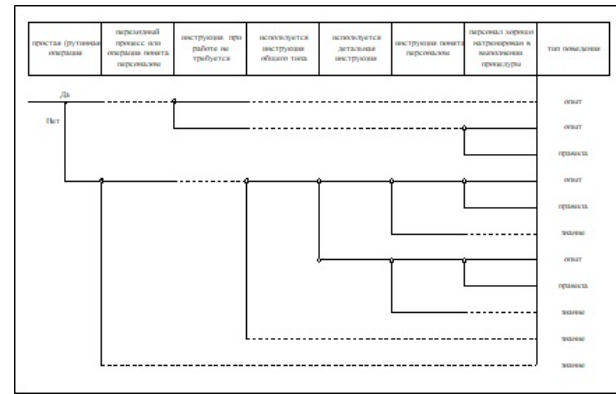


Рис. 1: Логическое дерево для определения преобладающего типа действий персонала при решении задачи

ставлены области вероятностей для описанных типов поведения [1, 2].

Логическая схема для помощи в решении к какому типу мыслительной деятельности относятся действия человека, показана на рис. 1.

Ошибки при восстановлении

Ошибки персонала, моделируемые в ВАБ по результатам восстановления, называются ошибками при восстановлении.

Процедуру восстановления можно разделить следующие этапы [3]:

- Обнаружение аварии.** Фаза обнаружения, как правило, не является проблемой. В момент, когда произошел останов реактора, срабатывают многочисленные сигнализации. Вероятность того, что операторы смены ничего не заметят, предельно мала и ею можно пренебречь.
- Определение причин текущего нарушения функционирования.** Самая чувствительная фаза. Вероятность ошибки во время этой фазы зависит от времени, которым располагает оператор и тип отказа, так как некоторые из них сложнее идентифицировать. Некорректная диагностика аварии влечет за собой ошибочные действия во время следующей фазы. За короткое время некорректная диагностика может

быть привести к совершению оператором действий, которые будут противоречить правильному курсу реального события, что, в конечном счете, лишь усугубит ситуацию.

3. Выбор процедуры, которую необходимо применить (если такая существует).

Процесс принятия решения оператором состоит из следующих этапов [3]:

1. Информационный поиск (таблица 2).

Таблица 2: Области вероятностей ошибок с учетом типов поведения

Действия	Операции	Включение психологических процессов
Обнаружение сигнала	Поисковое движение глаз, выделение сигнала из фона	Внимание, восприятие сигнала и фона
Сопоставление с эталоном	Сравнение по признакам	Наглядно-действенное мышление, оперативная память
Идентификация, опознание	Обнаружение сходства и различий, выделение существенных признаков, категоризация, запоминание	Оперативное мышление, долговременная память

2. Информационная подготовка принятия решения (таблица 3).

Таблица 3: Области вероятностей ошибок с учетом типов поведения

Действия	Операции	Включение психологических процессов
Переработка информации о проблемности ситуации	Операция построения образно-концептуальной модели (ОКМ) ситуации	Наглядно-действенное, наглядно-образное мышление
Поиск существенной информации. Выделение из фоновой, второстепенной информации. Классификация ситуации	Операции преобразования образов, сличения, обнаружения сходств и различий, выделение существенных признаков, категоризация.	Оперативная память. Долговременная память.
Обобщение информации. Построение ОКМ ситуации.	Операции по преобразованию ОКМ к виду, пригодному для принятия решения. Выбор оценочных критериев.	Оперативное мышление. Наглядно-образное мышление.

3. Принятие решения (таблица 4).

Таблица 4: Области вероятностей ошибок с учетом типов поведения

Действия	Операции	Включение психологических процессов
Определение цели	Операции формирования цели	Долговременная и оперативная память. Понятийно-образное мышление
Определение условий достижения цели	Операции формулирования условий достижения цели	Оперативное мышление. Теоретическое мышление
Разработка стратегии решения. Выбор оптимальной стратегии	Сопоставление разных методов решения	Долговременная память. Теоретическое мышление

R. Lazarus и R. Launier [5] определили психологический стресс как реакцию на особенности взаимодействия между личностью и окружающим миром. Это состояние — в большей степени продукт личностных когнитивных процессов, образа мыслей и оценки ситуации, знания собственных возможностей (ресурсов), степени обученности способам управления и стратегии поведения в экстремальных условиях, их адекват-

ному выбору. На оценку события как стрессогенного влияет ряд факторов, в т. ч. эмоции, ассоциирующиеся с данным событием, неопределенность ситуации, связанная с дефицитом информации для ее оценки, значимость события, отражающая степень его опасности для человека (или окружающих) и важность для достижения конечного результата.

Фазы физиологического процесса адаптации при стрессе

Начальная адаптация

Эта фаза развивается в самом начале действия (физиологического или патогенного факторов); складывается из двух разнонаправленных комплексов реакций [6]:

1. Ориентировочный рефлекс (сопровождается торможением многих видов осуществляющейся до этого времени деятельности).
2. Активизация нервно-трофического влияния (запасает и обеспечивает организм необходимой энергией).

Начальная фаза адаптации может быть выражена по-разному, в зависимости от силы раздражающих факторов (чем сильнее они, тем эта фаза более выражена). Соответственно, она может сопровождаться сильно или слабо выраженным эмоциональным компонентом.

Переходная фаза

В ходе этой фазы приспособительные механизмы организма постепенно переключаются на более глубокий тканевый уровень. Эти сдвиги обеспечивают новый уровень гомеостаза. Повышается неспецифическая резистентность организма и, в то же время, развиваются различные механизмы специфической адаптации [6].

Устойчивая адаптация

Носит долговременный характер. В основе предпосылок к развитию этой фазы лежат

механизмы памяти в ЦНС. Управляющие механизмы скоординированы. Характерной чертой жизнедеятельности в фазе устойчивой адаптации является относительная экономичность («выключение лишних реакций») затрат энергии для ее поддержания. Переключение реактивности организма на новый уровень сопровождается как мобилизацией ряда реакций, так и возвращением активности вспомогательных систем к исходным показателям [6].

Фаза дезадаптации

Это состояние может возникнуть в результате истощения физиологических резервов и нарушения взаимодействия регуляторных и метаболических механизмов адаптации. В результате в организме нарушается баланс расхода и восстановления в органах и тканях, а также взаимосвязь в работе физиологических систем. Вновь в состояние повышенной активности приходят вспомогательные системы — дыхания и кровообращения; энергия в организме тратится не экономно [6].

Дезадаптация возникает чаще всего в тех случаях, когда функциональная активность в новых условиях чрезмерна или действие факторов, явившихся основными стимуляторами адаптивных измерений в организме, усиливается, а они приближаются по силе к экстремальным.

Если требуются решительные действия для решения нестандартной задачи при дефиците времени, формирование целевой установки приобретает стрессовый характер. Чем противоречивее ситуативная и целевая установки, тем выше уровень испытываемого стресса.

Например, желание вывести управляемый агрегат из опасного состояния и осознание невозможности сделать это по объективным, ситуационным причинам или же когда работник не знает, как это в данной ситуации сделать. Это типичная ситуация, предшествующая вынужденному останову реактора, когда недостает времени или инфор-

мации на представление о том, что необходимо сделать для предотвращения останова.

Может иметь место и обратная картина — оператор знает, что делается в конкретных условиях угрозы выхода блока из заданного режима, но при наличии даже самого малого риска предпочитает предоставить работу автоматической системе защиты, которая остановит блок.

Заключение

В данной работе рассмотрены существующие методы проведения анализа надёжности персонала АЭС, описан процесс принятия решений персоналом. Также рассмотрены типы и различные виды ошибок персонала, а также влияющие на поведение факторы.

Таким образом, мы видим, что система адаптации при стрессе позволяет значительно снизить эмоциональную реакцию персонала. Именно эта система отвечает за выработку опыта, что способствует повышению работоспособности и корректности принятых решений. В то же время, пониженная эмоциональная реакция снижает внимательность, что приводит к совершению грубых ошибок оператором.

Литература

- [1] Сынгаевская А.А., Берберова М.А., Деревянкин А.А., Исламов Р.Т. Изучение влияния стресса на вероятность совершения ошибок персоналом АЭС в рамках анализа надёжности персонала / А.А.Сынгаевская, М.А.Берберова, А.А.Деревянкин, Р.Т.Исламов // СРТ1617: труды Международной научной конференции. — Протвино-Москва: Изд. ИФТИ, 2017, С. 339-342.
- [2] Сынгаевская А.А., Берберова М.А. Изучение механизма принятия решений персоналом АЭС в рамках анализа надёжности персонала / А.А. Сынгаевская, М.А. Берберова // SCVRT2017: труды Международной научной конференции. — Протвино-Москва: Изд. ИФТИ, 2017, с. 127-130.
- [3] Деревянкин, А.А. Исследование, разработка и применение методов оценки надёжности персонала при проведении вероятностного анализа безопасности атомных станций: диссертация на соискание ученой степени кандидата технических наук: 05.13.16 / Деревянкин Александр Альбертович. — М., 1990. — 200 с.

- [4] Rasmussen J. (1981) Models of Mental Strategies in Process Plant Diagnosis. In: Rasmussen J., Rouse W.B. (eds) Human Detection and Diagnosis of System Failures. NATO Conference Series, vol 15. Springer, Boston, MA.
- [5] Lazarus R. S., Launier R. Stress — related transactions between person and environment / L. A. Pervin, M. Lewis (Eds.), Perspectives in interactional psychology. New York: Pervin, 1978 pp. 287–327.
- [6] А.И.Яроцкий, И.А.Криволапчук. Эмоции человека в нормальных и стрессорных условиях. 2001.

Разработка атласа оценок показателей риска для российских АЭС*

*К.И.Чернявский*¹, *Н.К.Быков*¹, *С.С.Золотарев*¹, *М.А.Берберова*²

kirill.chernyavskiy@rambler.ru, maria.berberova@gmail.com

¹Московский физико-технический институт (государственный университет);

²Москва, Международный центр по ядерной безопасности

Согласно российскому законодательству (например, [1]), разработка паспорта безопасности объекта – обязательное условие функционирования таких объектов, как: гидротехнические сооружения, организации, производящие, перерабатывающие, хранящие или транспортирующие радиоактивные, пожаро- и взрывоопасные, токсичные химические и биологические вещества (АЭС, ТЭС, АЭС, промышленные предприятия и др.), а также характеризующиеся повышенным риском террористических действий (места скопления людей).

Ключевые слова: оценка риска, безопасность, ущерб, население, АЭС, атлас оценок показателей риска.

Systematization of available risk assessments for Russian nuclear power plants*

*K. I. Chernyavsky*¹, *N. K. Bykov*¹, *S. S. Zolotarev*¹, *M. A. Berberova*²

¹Moscow Institute of Physics and Technology (State University)

Moscow, ²International nuclear safety center

According to Russian legislation (for example, [1]), the development of an object safety certificate is an obligatory condition for the operation of such facilities as: hydraulic structures, organizations manufacturing, processing, storing or transporting radioactive, fire and explosive, toxic chemical and biological substances (gas stations, Thermal power plants, nuclear power plants, industrial enterprises, etc.), and also characterized by an increased risk of terrorist actions (places of congestion).

Keywords: Risk assessment, safety, damage, population, NPP, reactor, systematization, risk assessment atlas

Введение

Согласно ФЗ №184 от 27.12.02 г. «О техническом регулировании» [1, 2] риск определяется как вероятность причинения вреда жизни или здоровью граждан, имуществу физических или юридических лиц, государственному или муниципальному имуществу, окружающей среде, жизни или здоровью животных и растений с учетом тяжести этого вреда. Как уже было сказано, оценка таких вероятностных показателей производится, полученные данные учитываются при эксплуатации объектов, однако их систематизация и обобщение до сих пор не выполнялась.

За последние годы был проделан значительный объем работ в области ядерной безопасности, но население все еще относится к атомной энергетике с большим опасением, что может объясняться наличием

неопределенности при решении задач оценки рисков, а также пренебрежением рядом факторов, такими как состав населения, климатические условия и т.д. Кроме того, несмотря на крайне малую вероятность наиболее опасной аварии на атомной станции, зачастую последствия бывают существенные [2].

На данный момент известны два подхода к оценке риска: детерминистский и вероятностный [2]. Целью детерминистского подхода является обоснование исполнения полученных критериев безопасности при помощи предусмотренных технических средств и организационных мер при эксплуатации и при любом из учтенных исходных событий с учетом принципа единичного отката.

При расчетах в рамках вероятностного анализа безопасности (ВАБ) учитываются реалистичные предпосылки в части исходных состояний и вариантов дальнейшего развития последствий аварий. В рамках ВАБ рассматриваются следующие понятия:

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 16-07-00089, 18-07-20030, 18-37-10008.

- Проектная (наиболее вероятная) авария (ПА) – авария, для которой проектом определены исходные события, конечные состояния, системы безопасности;
- Запроектная (наиболее опасная) авария (ЗПА) – авария, которая развивалась из исходного события, для которого не предусмотрена система безопасности из-за малой вероятности данного события или невозможности иметь разумные меры по защите от него.

Основной проблемой ВАБ является затратный, с точки зрения трудовых ресурсов, сбор информации об изменении состояния системы и её параметров при многочисленных состояниях. К примеру, затруднительно получить данные по отказам системы безопасности, на основании которых рассчитывается вероятность ЗПА.

Из всего вышесказанного следует, что на текущий момент степень неопределенности исходных данных и отсутствие систематизированного учета ряда дополнительных факторов при использовании рассмотренных подходов оказывает большое влияние на оценку риска аварий на АЭС. В связи с этим в данной работе ставится задача систематизации оценки рисков для различных АЭС с учетом ряда дополнительных факторов, таких как состав населения с учетом возрастного критерия и климатический фактор, а именно наиболее частое направление ветра на месте расположения АЭС [2].

В работе авторами предпринята попытка формирования атласа оценок показателей риска по всем энергоблокам Российских АЭС.

1. Расчет доз внешнего и внутреннего облучения населения в кольцевом сегменте румба

Расчет годовой эффективной дозы облучения населения в кольцевом сегменте румба, с учетом розы ветров, проводился по следующей формуле [2, 3, 4].

При равномерном распределении населения в кольцевом сегменте румба

$$D_E(R) = \frac{P_W}{R} \int_0^R N_{1/8}(r) D_{E(year)}(r) dr, \quad (1)$$

где:

D_E – годовая эффективная доза облучения (effective dose) населения в кольцевом сегменте румба, с учетом розы ветров;

R, r – расстояние от АЭС;

P_W – вероятность направления ветра (probability of the wind direction);

$N_{1/8}$ – количество людей в румбе;

$D_{E(year)}$ – годовая эффективная доза облучения всего населения.

При дискретном распределении населения в кольцевом сегменте румба

$$D_E(R) = P_W \sum_{i,k} N_{1/8(i)} D_{E(year)(i)}, \quad (2)$$

где:

D_E – годовая эффективная доза облучения (effective dose) населения в кольцевом сегменте румба, с учетом розы ветров;

R – расстояние от АЭС;

P_W – вероятность направления ветра (probability of the wind direction);

$N_{1/8}$ – количество людей в румбе;

$D_{E(year)}$ – годовая эффективная доза облучения всего населения.

2. Расчет ущерба населению в кольцевом сегменте румба, в результате воздействия радиоактивных веществ

Расчет ущерба населению в кольцевом сегменте румба, в результате воздействия радиоактивных веществ, проводился по следующей формуле [2, 3, 4].

При равномерном распределении населения в кольцевом сегменте румба:

$$L(R) = \frac{L_\Sigma}{N_\Sigma} \cdot \frac{\int_0^R N_{1/8}(r) D_E(r) dr}{D_{E(year)}}, \quad (3)$$

где:

L – ущерб, нанесенный населению в кольцевом сегменте румба;

R, r – расстояние от АЭС;

L_{Σ} – общий ущерб от аварии;

N_{Σ} – общая численность населения;

D_E – годовая эффективная доза облучения населения в кольцевом сегменте румба, с учетом розы ветров;

$N_{1/8}$ – количество людей в румбе;

$D_{E(year)}$ – годовая эффективная доза облучения всего населения.

При дискретном распределении населения в кольцевом сегменте румба

$$L(R) = \frac{L_{\Sigma}}{N_{\Sigma}} \cdot \frac{\sum_{i,k} N_{1/8(i)}(R) D_{E(i)}(R)}{D_{E(year)}} \quad (4)$$

где:

L – ущерб, нанесенный населению в кольцевом сегменте румба;

R – расстояние от АЭС;

L_{Σ} – общий ущерб от аварии;

N_{Σ} – общая численность населения;

D_E – годовая эффективная доза облучения населения в кольцевом сегменте румба, с учетом розы ветров;

$N_{1/8}$ – количество людей в румбе;

$D_{E(year)}$ – годовая эффективная доза облучения всего населения;

i, k – расстояние от АЭС.

Результаты расчетов по всем российским АЭС приведены в работах [2] - [20].

Проанализируем результаты расчетов по формуле (4). Получаем, что на расстоянии 3-100 км ущерб, нанесенный одному человеку, не превышает 50 руб. В таблице 1 [2] приведены оценки (вновь рассчитанные и уточненные) доз максимального облучения одного человека и ущерба, нанесенного одному человеку, в кольцевом сегменте румба

на расстоянии 3-100 км от АЭС для Белоярской, Билибинской, Калининской, Курской, Ростовской, Смоленской, Балаковской и Нововоронежской АЭС.

Таблица 1: Оценки максимального облучения и ущерба на расстоянии 3-100 км от АЭС

АЭС	Максимальное облучение, мЗв	Максимальный ущерб, нанесенный одному человеку, руб.
Белоярская	8	285,5
Билибинская	5,22	130,4
Калининская	1,1	74,5
Курская	3	15
Ростовская	2,5	16,67
Смоленская	1,6	65,40
Балаковская	0,8	18,35
Кольская	1,4	48,54
Нововоронежская	0,6	67

3. Формирование атласа оценок показателей риска энергоблоков российских АЭС

На сегодняшний день формирование атласа оценок показателей риска проводится по работам [2] - [20].

В данной статье предлагается пример формата атласа оценок показателей риска.

Для энергоблоков с реакторами типа РБМК-1000 (Курская и Смоленская АЭС) – таблицы 2-13, рисунки 1-6.

3.1. Курская АЭС

На рис. 1 представлена диаграмма повторяемости направления ветра на Курской АЭС [3].

На рис. 2 приведено графическое представление результатов расчета годовых эффективных доз облучения одного человека и всего населения, проживающего в кольцевом сегменте румба, с учетом розы ветров, для Курской АЭС [3].

По оси X представлено распределение годовых эффективных доз облучения одного человека с учетом розы ветров (Зв). Принято, что:

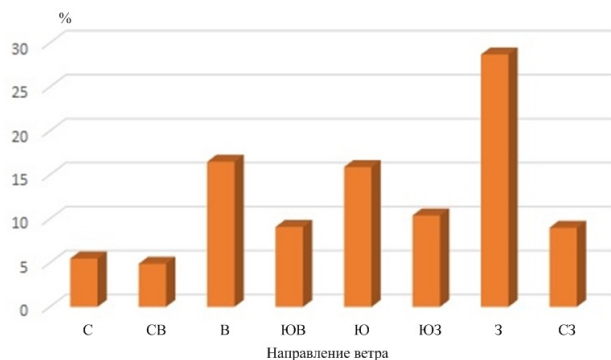


Рис. 1: Диаграмма повторяемости направления ветра на Курской АЭС

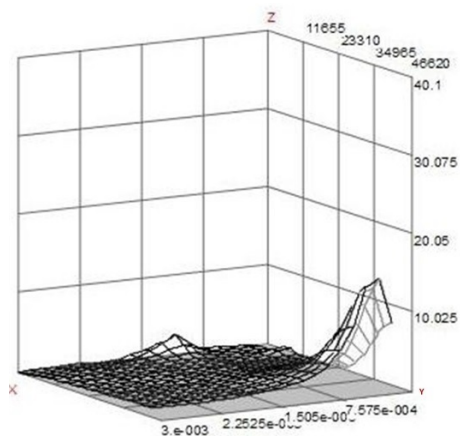


Рис. 2: Графическое представление результатов расчета годовых эффективных доз облучения одного человека и всего населения, проживающего в кольцевом сегменте румба, с учетом розы ветров, для Курской АЭС

- минимальная годовая эффективная доза облучения одного человека с учетом розы ветров (X_{min}) составляет 10^{-5} Зв;
- максимальная годовая эффективная доза облучения одного человека с учетом розы ветров (X_{max}) составляет $3 \cdot 10^{-3}$ Зв.

По оси Y представлено количество людей в кольцевом сегменте румба (чел.). Принято, что:

- минимальное количество людей в кольцевом сегменте румба (Y_{min}) составляет 0 чел.;
- максимальное количество людей в кольцевом сегменте румба (Y_{max}) составляет 46620 чел.

По оси Z представлено распределение коллективных доз облучения населения в коль-

цевом сегменте румба (чел·Зв). Принято, что:

- минимальная коллективная доза облучения населения в кольцевом сегменте румба (Z_{min}) составляет 0 чел·Зв;
- максимальная коллективная доза облучения населения в кольцевом сегменте румба (Z_{max}) составляет 40.1 чел·Зв.

На рис. 3 приведено графическое представление результатов расчета ущерба, нанесенного одному человеку и всему населению, проживающему в кольцевом сегменте румба, на расстоянии 3-30 км от АЭС для Курской АЭС [3].

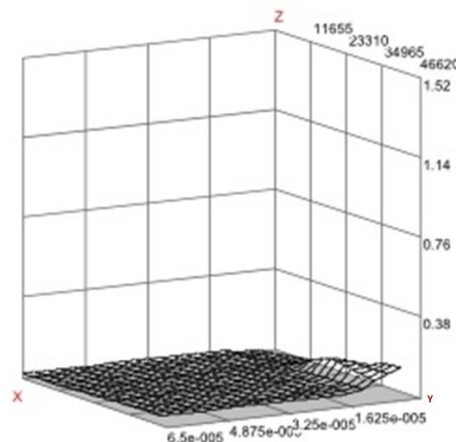


Рис. 3: Графическое представление результатов расчета ущерба, нанесенного одному человеку и всему населению, проживающему в кольцевом сегменте румба, на расстоянии 3-30 км от АЭС для Курской АЭС

По оси X представлено распределение годовых эффективных доз облучения одного человека с учетом розы ветров (Зв). Принято, что:

- минимальный ущерб, нанесенный одному человеку в кольцевом сегменте румба (X_{min}) составляет 0 руб.;
- максимальный ущерб, нанесенный одному человеку в кольцевом сегменте румба (X_{max}) составляет $6.5 \cdot 10^{-5}$ млн.руб.

По оси Y представлено количество людей в кольцевом сегменте румба (чел.). Принято, что:

- минимальное количество людей в кольцевом сегменте румба (Y_{min}) составляет 0 чел.;
- максимальное количество людей в кольцевом сегменте румба (Y_{max}) составляет 46620 чел.

По оси Z представлено распределение ущерба, нанесенного всему населению в кольцевом сегменте румба (млн. руб.). Принято, что:

- минимальный ущерб, нанесенный всему населению в кольцевом сегменте румба (Z_{min}) составляет 0 руб.;
- максимальный ущерб, нанесенный всему населению в кольцевом сегменте румба (Z_{max}) составляет 1.52 млн.руб.

Таблица 2: Показатели радиационного риска для населения от аварий на Курской АЭС [5]

Средняя индивидуальная эффективная доза для населения E_i , Зв	Индивидуальный риск смерти населения (отдаленные эффекты) r_c , (на 1 аварию)	Индивидуальный риск смерти населения r_{ci} , 1/год (с учетом частоты ЧС)
$2.5 \cdot 10^{-4}$	$1.25 \cdot 10^{-5}$	$8.75 \cdot 10^{-13}$

Таблица 3: Показатели экономического риска для населения от аварий на Курской АЭС [5]

Численность населения, чел	Коллективная доза для населения, чел·Зв	Ущерб для населения, млн. руб.
1 365 132	339	102

Таблица 4: Вероятность направления распространения радиоактивных веществ по сторонам света на Курской АЭС [5]

Румб	Повторяемость направления ветра, %
С	5.5
СВ	4.9
В	16.5
ЮВ	9.1
Ю	15.9
ЮЗ	10.4
З	28.7
СЗ	9.0
Всего	100

Таблица 6: Результаты расчета годовых эффективных доз облучения одного человека и всего населения, проживающего в кольцевом сегменте румба, с учетом розы ветров, на расстоянии 3-30 км от АЭС для Курской АЭС [3, 4]

Румб	Годовая эффективная доза облучения одного человека с учетом розы ветров (Зв)	Кол-во людей, чел.	Коллективная доза облучения населения в кольцевом сегменте румба (чел·Зв)
Расстояние 3-10 км			
Ю	$2.3 \cdot 10^{-4}$	3346	0.76
ЮЗ	$2 \cdot 10^{-4}$	10552	2.14
З	$6.8 \cdot 10^{-4}$	435	0.30
СЗ	$3.8 \cdot 10^{-4}$	809	0.30
С	$6.6 \cdot 10^{-4}$	691	0.45
СВ	$4.3 \cdot 10^{-4}$	636	0.27
В	$1.2 \cdot 10^{-3}$	2471	2.93
ЮВ	$3.7 \cdot 10^{-4}$	46620	17.34
Расстояние 10-15 км			
Ю	$3 \cdot 10^{-5}$	321	0.01
ЮЗ	$2 \cdot 10^{-5}$	371	0.01
З	$8 \cdot 10^{-5}$	330	0.02
СЗ	$4 \cdot 10^{-5}$	238	0.01
С	$7 \cdot 10^{-5}$	184	0.01
СВ	$5 \cdot 10^{-5}$	600	0.03
В	$1.3 \cdot 10^{-4}$	941	0.12
ЮВ	$4 \cdot 10^{-5}$	649	0.03
Расстояние 15-20 км			
Ю	10^{-5}	1260	0.02
ЮЗ	10^{-5}	2010	0.02
З	$4 \cdot 10^{-5}$	1303	0.05
СЗ	$2 \cdot 10^{-5}$	1808	0.04
С	$4 \cdot 10^{-5}$	755	0.03
СВ	$3 \cdot 10^{-5}$	2303	0.06
В	$7 \cdot 10^{-5}$	1953	0.14
ЮВ	$2 \cdot 10^{-5}$	3114	0.07
Расстояние 20-30 км			
Ю	10^{-5}	1352	0.01
ЮЗ	10^{-5}	6900	0.05
З	$2 \cdot 10^{-5}$	29779	0.72
СЗ	10^{-5}	12345	0.17
С	$2 \cdot 10^{-5}$	1661	0.04
СВ	$2 \cdot 10^{-5}$	4853	0.07
В	$4 \cdot 10^{-5}$	20631	0.87
ЮВ	10^{-5}	5695	0.08

Таблица 5: Распределение населения на расстоянии 3-100 км от Курской АЭС [5]

Румб	3-5	5-10	10-15	15-20	20-30	30-40	40-50	50-100
С	467	224	184	755	1 661	7 700	9 880	82 432
СВ	0	636	600	2 303	4 853	7 700	9 880	82 432
В	0	2471	941	1 953	20631	7 700	408000	82 432
ЮВ	46500	120	649	3 114	5 695	7 700	9 880	82 432
Ю	0	3346	321	1 260	1 352	7 700	9 880	82 432
ЮЗ	0	10552	371	2 010	6 900	7 700	9 880	82 432
З	0	435	330	1 303	29779	7 700	9 880	82 432
СЗ	685	124	238	1 808	12345	7 700	9 880	82 432
Всего	47652	17908	3634	14506	83216	61600	477160	659456

Таблица 7: Результаты расчета ущерба одному человеку, проживающему в кольцевом сегменте румба, на расстоянии 3-30 км от АЭС для Курской АЭС [3]

Румб	Ущерб, нанесенный одному человеку в кольцевом сегменте румба, руб.	Кол-во людей в румбе, чел.
Расстояние 3-10 км		
Ю	2.7	3346
ЮЗ	2.4	10552
З	8.2	435
СЗ	4.5	809
С	7.9	691
СВ	5.2	636
В	14	2471
ЮВ	4.5	46620
Расстояние 10-15 км		
Ю	2.7	321
ЮЗ	2.4	371
З	8.2	330
СЗ	4.5	238
С	7.9	184
СВ	5.2	600
В	14	941
ЮВ	4.5	649
Расстояние 15-20 км		
Ю	2.8	1260
ЮЗ	2.4	2010
З	8.3	1303
СЗ	4.5	1808
С	7.9	755
СВ	5.1	2303
В	14	1953
ЮВ	4.5	3114
Расстояние 20-30 км		
Ю	2.7	1352
ЮЗ	2.4	6900
З	8.2	29779
СЗ	4.4	12345
С	7.8	1661
СВ	5.1	4853
В	0.1	20631
ЮВ	4.4	5695

3.2. Смоленская АЭС

На рис. 4 представлена диаграмма повторяемости направления ветра на Смоленской АЭС [3].

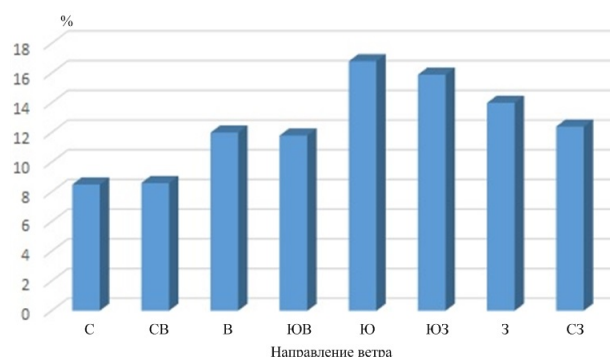


Рис. 4: Диаграмма повторяемости направления ветра на Смоленской АЭС

На рис. 5 приведено графическое представление результатов расчета годовых эффективных доз облучения одного человека и всего населения, проживающего в кольцевом сегменте румба, с учетом розы ветров, для второй очереди Смоленской АЭС [3].

По оси X представлено распределение годовых эффективных доз облучения одного человека с учетом розы ветров ($Zв$). Принято, что:

- минимальная годовая эффективная доза облучения одного человека с учетом розы ветров (X_{min}) составляет 10^{-5} Зв;
- максимальная годовая эффективная доза облучения одного человека с учетом розы ветров (X_{max}) составляет $3 \cdot 10^{-3}$ Зв.

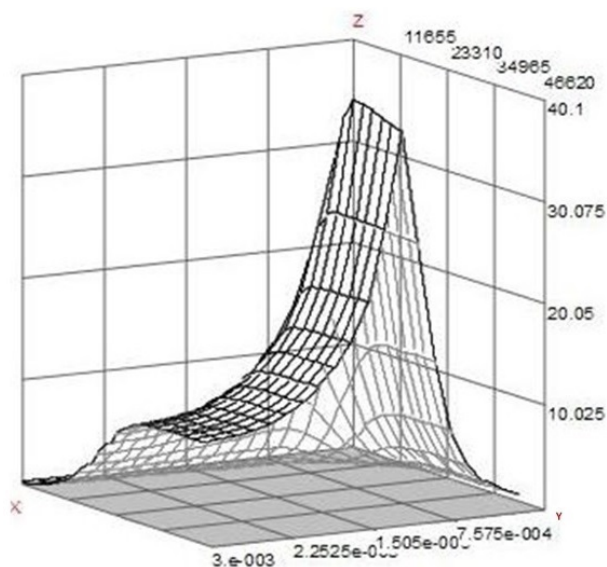


Рис. 5: Графическое представление результатов расчета годовых эффективных доз облучения одного человека и всего населения, проживающего в кольцевом сегменте румба, с учетом розы ветров, для Смоленской АЭС

По оси Y представлено количество людей в кольцевом сегменте румба (чел.). Принято, что:

- минимальное количество людей в кольцевом сегменте румба (Y_{min}) составляет 0 чел.;
- максимальное количество людей в кольцевом сегменте румба (Y_{max}) составляет 46620 чел.

По оси Z представлено распределение коллективных доз облучения населения в кольцевом сегменте румба (чел·Зв). Принято, что:

- минимальная коллективная доза облучения населения в кольцевом сегменте румба (Z_{min}) составляет 0 чел·Зв;
- максимальная коллективная доза облучения населения в кольцевом сегменте румба (Z_{max}) составляет 40.1 чел·Зв.

На рис. 6 приведено графическое представление результатов расчета ущерба, нанесенного одному человеку и всему населению, проживающему в кольцевом сегменте румба, на расстоянии 3-30 км от АЭС для Смоленской АЭС [3].

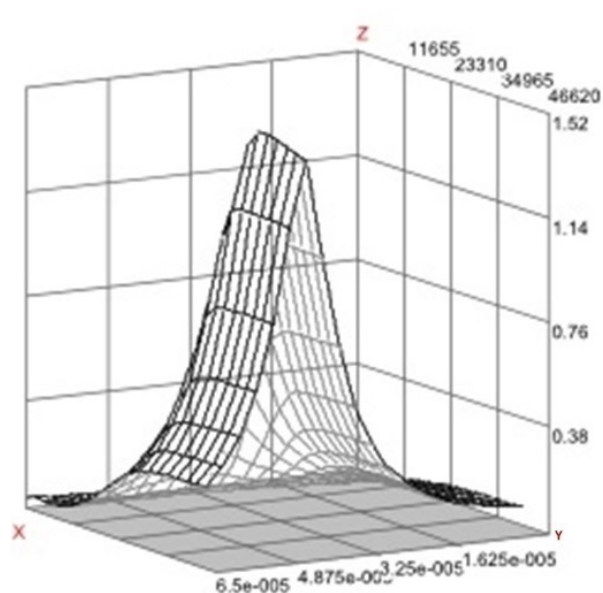


Рис. 6: Графическое представление результатов расчета ущерба, нанесенного одному человеку и всему населению, проживающему в кольцевом сегменте румба, на расстоянии 3-30 км от АЭС для Смоленской АЭС

По оси X представлено распределение годовых эффективных доз облучения одного человека с учетом розы ветров (Зв). Принято, что:

- минимальный ущерб, нанесенный одному человеку в кольцевом сегменте румба (X_{min}) составляет 0 руб.;
- максимальный ущерб, нанесенный одному человеку в кольцевом сегменте румба (X_{max}) составляет $6.5 \cdot 10^{-5}$ млн.руб.

По оси Y представлено количество людей в кольцевом сегменте румба (чел.). Принято, что:

- минимальное количество людей в кольцевом сегменте румба (Y_{min}) составляет 0 чел.;
- максимальное количество людей в кольцевом сегменте румба (Y_{max}) составляет 46620 чел.

По оси Z представлено распределение ущерба, нанесенного всему населению в кольцевом сегменте румба (млн. руб.). Принято, что:

- минимальный ущерб, нанесенный всему населению в кольцевом сегменте румба (Z_{min}) составляет 0 руб.;

— максимальный ущерб, нанесенный всему населению в кольцевом сегменте румба (Z_{max}) составляет 1.52 млн.руб.

Таблица 8: Показатели радиационного риска от аварий на Смоленской АЭС [6]

Средняя индивидуальная эффективная доза для населения E_i , Зв	Индивидуальный риск смерти населения (отдаленные эффекты) r_c , (на 1 аварию)	Индивидуальный риск смерти населения r_{c_i} , 1/год (с учетом частоты ЧС)
$1.3 \cdot 10^{-3}$	$6.4 \cdot 10^{-5}$	$4.5 \cdot 10^{-12}$

Таблица 9: Показатели экономического риска для населения от аварий на Смоленской АЭС [6]

Численность населения, чел	Коллективная доза для населения, чел·Зв	Ущерб для населения, млн. руб.
257 412	330	99

Таблица 10: Вероятность направления распространения радиоактивных веществ по сторонам света на Смоленской АЭС [6]

Румб	Повторяемость направления ветра, %
С	8.5
СВ	8.6
В	12
ЮВ	11.8
Ю	16.8
ЮЗ	15.9
З	14
СЗ	12.4
Всего	100

Таблица 12: Результаты расчета годовых эффективных доз облучения одного человека и всего населения, проживающего в кольцевом сегменте румба, с учетом розы ветров, на расстоянии 3-30 км от АЭС для Смоленской АЭС [3]

Румб	Годовая эффективная доза облучения одного человека с учетом розы ветров (Зв)	Кол-во людей, чел.	Коллективная доза облучения населения в кольцевом сегменте румба (чел·Зв)
Расстояние 3-10 км			
Ю	$8.3 \cdot 10^{-4}$	0	0
ЮЗ	$8.4 \cdot 10^{-4}$	0	0
З	$1.2 \cdot 10^{-3}$	0	0
СЗ	$1.2 \cdot 10^{-3}$	0	0
С	$1.6 \cdot 10^{-3}$	0	0
СВ	$1.6 \cdot 10^{-3}$	0	0
В	$1.4 \cdot 10^{-3}$	0	0
ЮВ	$1.2 \cdot 10^{-3}$	33000	40.1
Расстояние 10-15 км			
Ю	$9.4 \cdot 10^{-5}$	294	0.03
ЮЗ	$9.5 \cdot 10^{-5}$	294	0.03
З	$1.3 \cdot 10^{-4}$	294	0.04
СЗ	$1.3 \cdot 10^{-4}$	294	0.04
С	$1.8 \cdot 10^{-4}$	294	0.05
СВ	$1.7 \cdot 10^{-4}$	294	0.05
В	$1.5 \cdot 10^{-4}$	294	0.05
ЮВ	$1.4 \cdot 10^{-4}$	294	0.04
Расстояние 15-20 км			
Ю	$5 \cdot 10^{-5}$	412	0.02
ЮЗ	$5.1 \cdot 10^{-5}$	412	0.02
З	$7.1 \cdot 10^{-5}$	412	0.03
СЗ	$7 \cdot 10^{-5}$	412	0.03
С	$9.9 \cdot 10^{-5}$	412	0.04
СВ	$9.4 \cdot 10^{-5}$	412	0.04
В	$8.3 \cdot 10^{-5}$	412	0.03
ЮВ	$7.3 \cdot 10^{-5}$	412	0.03
Расстояние 20-30 км			
Ю	$3 \cdot 10^{-5}$	1176	0.04
ЮЗ	$3 \cdot 10^{-5}$	1176	0.04
З	$4.2 \cdot 10^{-5}$	1176	0.05
СЗ	$4.1 \cdot 10^{-5}$	1176	0.05
С	$5.9 \cdot 10^{-5}$	1176	0.07
СВ	$5.6 \cdot 10^{-5}$	1176	0.07
В	$4.9 \cdot 10^{-5}$	1176	0.06
ЮВ	$4.3 \cdot 10^{-5}$	1176	0.05

Таблица 11: Распределение населения на расстоянии 3-100 км от Смоленской АЭС [6]

Румб	3-5	5-10	10-15	15-20	20-30	30-40	40-50	50-100
С	0	0	294	412	1176	1512	1942	16192
СВ	0	0	294	412	1176	1512	1942	16192
В	0	0	294	412	1176	1512	1942	16192
ЮВ	31000	2000	294	412	1176	1512	1942	16192
Ю	0	0	294	412	1176	1512	1942	16192
ЮЗ	0	0	294	412	1176	53700	1942	16192
З	0	0	294	412	1176	1512	1942	16192
СЗ	0	0	294	412	1176	1512	1942	16192
Всего	31000	2000	2352	3296	9408	64284	15536	129536

Таблица 13: Результаты расчета ущерба одному человеку, проживающему в кольцевом сегменте румба, на расстоянии 3-30 км от АЭС для Смоленской АЭС [3]

Румб	Кол-во людей в румбе, чел.	Ущерб, нанесенный одному человеку в кольцевом сегменте румба, руб.
Расстояние 3-10 км		
Ю	0	0
ЮЗ	0	0
З	0	0
СЗ	0	0
С	0	0
СВ	0	0
В	0	0
ЮВ	33000	40.1
Расстояние 10-15 км		
Ю	294	0.03
ЮЗ	294	0.03
З	294	0.04
СЗ	294	0.04
С	294	0.05
СВ	294	0.05
В	294	0.05
ЮВ	294	0.04
Расстояние 15-20 км		
Ю	412	0.02
ЮЗ	412	0.02
З	412	0.03
СЗ	412	0.03
С	412	0.04
СВ	412	0.04
В	412	0.03
ЮВ	412	0.03
Расстояние 20-30 км		
Ю	1176	0.04
ЮЗ	1176	0.04
З	1176	0.05
СЗ	1176	0.05
С	1176	0.07
СВ	1176	0.07
В	1176	0.06
ЮВ	1176	0.05

4. Основные результаты и выводы к работе

На данном этапе работы было выполнено:

1. Рассмотрены материалы по оценкам риска энергоблоков всех российских АЭС.
2. Предложен вариант формата атласа оценок риска.
3. Выдвинуты первоначальные предположения относительно расхождения в полученных результатах как для доз, так и для ущерба. Предположено, что одной из возможных причин расхождения в результатах расчета дозы и ущерба для Курской и Смоленской АЭС может быть разница в рельефе местности. Нам необходимо это учесть. Под изменениями рельефа будем понимать не овраги и склоны, а холмы, горы, поля. Дополнительно можно учитывать и тип местности: степи, леса и прочее, хотя все это будет вносить свой вклад на уже достаточно большом отдалении от АЭС.

В дальнейшем планируется продолжить работу над атласом оценок риска и продумать более удобный его формат.

Литература

- [1] О техническом регулировании: [Федеральный закон №184-ФЗ от 27.12.02: принят Государственной Думой Российской Федерации 15 декабря 2002 года]. – М., 2002. – 84 с.
- [2] Быков Н.К., Золотарев С.С., Чернявский К.И., Берберова М.А. Разработка атласа оценок показателей риска для российских АЭС / НК. Быков, С.С. Золотарев, К.И. Чернявский, М.А. Берберова // SCVRT2017: труды Международной научной конференции. – Протвино-Москва: Изд. ИФТИ. – С. 85-89.
- [3] Берберова М.А. Оценка показателей риска для вторых очередей Смоленской и Курской АЭС: диссертация на соискание ученой степени кандидата технических наук: 05.14.03 / Берберова Мария Александровна. – М., 2015. – 130 с.

- [4] Берберова М.А., Исламов Р.Т., Зарипов И.Р., Липатов А.А., Лебедев К.Ю., Гриднев А.А., Бурханов Р.А., Воцан И.О. Оценка показателей риска для АЭС с реакторами типа БН-600 и РБМК-1000 // Труды Международной конференции и Школы по стойкости социотехнических систем Resilience2014, 25-28 ноября 2014 г., Протвино, Парк Дракино, Московская область. — Изд. ИФТИ, Протвино-Москва, 2015, ISBN 978-5-88835-037-9, С.179-185.
- [5] Корректировка паспорта безопасности Курской АЭС. Оценка показателей риска Курской АЭС для разработки раздела II Паспорта безопасности: отчет о НИР рег. № 2152ОТ11 / Кабанов Л.П., Берберова М.А. [и др.] — М.: Международный Центр по Ядерной Безопасности, 2011. — 68 с.
- [6] Разработка перечня мероприятий по управлению безопасностью и оценка показателей риска для 2 очереди (3 блока) Смоленской АЭС с реактором РБМК-1000: отчет о НИР рег. № 2314ОТ12 / Кабанов Л.П., Берберова М.А. [и др.] — М.: Международный Центр по Ядерной Безопасности, 2012. — 225 с.
- [7] Берберова М.А., Исламов Р.Т., Липатов А.А., Орлов Ю.Н. Развитие методического подхода к оценке доз облучения / Труды Международной научной конференции СРТ2015. ИФТИ, 2016. С. 260-263.
- [8] Липатов А.А. Оценка показателей риска для населения, проживающего вокруг АЭС (на примере Билибинской АЭС): диссертация на соискание ученой степени магистра: 010900 / Липатов Александр Андреевич. — М., 2017. — 28 с.
- [9] Максимов А.С., Берберова М.А. Сравнительная оценка доз внешнего и внутреннего облучения и оценка ущерба населению, проживающему вокруг АЭС с реакторами типа ВВЭР (на примере Ростовской и Калининской АЭС) / А. С. Максимов, М. А. Берберова // СРТ1617: труды Международной научной конференции. — Протвино-Москва: Изд. ИФТИ, 2017, С. 257-260.
- [10] Берберова М.А., Исламов Р.Т., Лебедев К.Ю. Оценка доз внешнего и внутреннего облучения населения, проживающего вокруг АЭС (на примере Ростовской АЭС) / М. А. Берберова, Р.Т. Исламов, К.Ю.Лебедев // СРТ-2015: труды Международной научной конференции. — Протвино-Москва: Изд. ИФТИ. — С. 256-259.
- [11] Берберова М.А., Исламов Р.Т., Гриднев А.А., Лебедев К.Ю. Оценка ущерба населению, проживающему вокруг АЭС, в результате воздействия радиоактивных веществ / М.А. Берберова, Р.Т. Исламов, А.А. Гриднев, К.Ю. Лебедев // СРТ-2015: труды Международной научной конференции. — Протвино-Москва: Изд. ИФТИ. — С. 230-232.
- [12] Сапаркин Л.М., Берберова М.А. Разработка модели оценки ущерба, нанесенного населению в результате воздействия радиоактивных веществ при аварии на АЭС (на примере Нововоронежской АЭС) / М.А. Берберова, Л.М. Сапаркин // SCVRT2017: труды Международной научной конференции. — Протвино-Москва: Изд. ИФТИ. — С. 175-183.
- [13] Разработка перечня мероприятий по управлению безопасностью и оценка показателей риска для блока Балаковской АЭС с реактором ВВЭР-1000: отчет о НИР рег. № 2031ОТ11 / Кабанов Л.П., Берберова М.А. [и др.] — М.: Международный Центр по Ядерной Безопасности, 2011. — 73 с.
- [14] Разработка перечня мероприятий по управлению безопасностью и оценка показателей риска для АЭС с реакторами типа БН (Белоярская АЭС с реактором БН-600): отчет о НИР рег. № 2155ОТ11 / Кабанов Л.П., Берберова М.А. [и др.] — М.: Международный Центр по Ядерной Безопасности, 2011. — 73 с.
- [15] Методика оценки показателей риска АЭС для паспортов безопасности и управления безопасностью Билибинской АЭС с реакторами типа ЭГП-6: отчет о НИР рег. № 2398ОТ13 / Берберова М.А. [и др.] — М.: Международный Центр по Ядерной Безопасности, 2013. — 129 с.
- [16] Разработка перечня мероприятий по управлению безопасностью и оценка показателей риска для 2 очереди Калининской АЭС с реактором ВВЭР-1000 (В-320): отчет о НИР рег. № 2339ОТ12 / Берберова М.А. [и др.] — М.: Международный Центр по Ядерной Безопасности, 2012. — 109 с.
- [17] Разработка перечня мероприятий (рекомендаций) по управлению безопасностью, оценка показателей риска для энергоблоков Кольской АЭС с реактором ВВЭР-440 (В-213): отчет о НИР рег. № 2352ОТ13 / Берберова М.А. [и др.] — М.: Международный Центр по Ядерной Безопасности, 2013. — 114 с.
- [18] Разработка перечня мероприятий по управлению безопасностью и оценка показателей риска для энергоблоков Нововоронежской АЭС с реактором ВВЭР-440 (В-179/230): отчет о НИР рег. № 2376ОТ13 / Берберова М.А. [и др.] — М.: Международный Центр по Ядерной Безопасности, 2013. — 80 с.
- [19] Разработка перечня мероприятий по управлению безопасностью и оценка показателей риска для энергоблоков Нововоронежской АЭС с реактором ВВЭР-1000: отчет о НИР рег. № 2044ОТ11 / Кабанов Л.П., Берберова М.А. [и др.] — М.: Международный Центр по Ядерной Безопасности, 2011. — 77 с.
- [20] Оценка риска референтной АЭС с ВВЭР (действующий блок АЭС с ВВЭР-1000): отчет о НИР рег. № 1821ОТ10 / Кабанов Л.П., Берберова М.А. [и др.] — М.: Международный Центр по Ядерной Безопасности, 2010. — 129 с.

Геометрическое моделирование и оптимизация многофакторных процессов и явлений методом многомерной интерполяции*

Е. В. Конопацкий

e.v.konopatskiy@mail.ru

Макеевка, Донбасская национальная академия строительства и архитектуры

В статье рассматриваются вопросы геометрического моделирования многофакторных процессов и явлений с помощью многомерной интерполяции реализованной в БН-исчислении. Приведены преимущества и обоснован выбор БН-исчисления как аппарата геометрического моделирования многофакторных процессов и явлений. Представлена общая концепция метода многомерной интерполяции, состоящая из 7 этапов. Также показаны этапы формирования дерева геометрической модели. Аналитическое описание геометрической модели представлено в виде последовательности точечных уравнений. Предложено описание способа конструирования дуг алгебраических кривых проходящих через наперёд заданные точки.

Ключевые слова: геометрическое моделирование, многомерная интерполяция, многофакторный процесс, геометрический объект, функция отклика.

Geometrical modeling and optimization of multifactor processes and phenomena's by method of multidimensional interpolation*

E. V. Konopatskiy

Makeyevka, Donbas National Academy of Civil Engineering and Architecture

The article deals with the problems of geometric modeling of multifactor processes and a phenomena's using multidimensional interpolation implemented in BN-calculation. The advantages and justification of the choice of BN-calculation as an apparatus for geometric modeling of multifactor processes and phenomena's are presented. The general concept of the multidimensional interpolation method consisting of 7 stages is presented. The stages of forming a tree of a geometric model are also shown. Analytical description of the geometric model is represented as a sequence of point equations. A method for constructing arcs of algebraic curves passing through given points is proposed.

Keywords: geometric modeling, multidimensional interpolation, multifactor process, geometric object, response function.

Введение

Задачи многофакторных процессов и явлений являются неотъемлемой частью любой области знаний, накопленных человеческим обществом, что обеспечивает широкий спектр инструментов и методов их решения. Тем не менее, существующие методы моделирования и оптимизации не всегда могут обеспечить полный учет функциональных, конструктивных, технологических, экономических, эстетических и других требований, необходимых для решения научных и прикладных задач моделирования и оптимизации, поскольку выполняют анализ факторов, влияющих на протекание процесса (или явления) по очереди, что не даёт возможность оценить влияние того или иного фактора на весь про-

цесс в целом. На данный момент, эффективное решение подобного класса задач является возможным с применением методов многомерной геометрии, которые позволяют представить любой многофакторный процесс в виде многопараметрического геометрического объекта, принадлежащего многомерному аффинному пространству, с последующей его оптимизацией высокоточными методами математического анализа. Особенно эффективно применение метода многомерной интерполяции при решении задач, для которых проведение эксперимента с реальной системой, как минимум, не рентабельно, а иногда просто и не возможно. Однако аналитическое описание геометрических объектов многомерного пространства затрудняется из-за ограниченного инструментария существующих методов моделирования, для которых, учитывая сложность визуального восприятия многомер-

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20030.

ного пространства, на первую роль выходит не зрительная наглядность, а логическая, основанная на методах обобщения и аналогии. Поэтому разработку теоретических основ и практических методов геометрического моделирования многофакторных процессов и явлений, основанных на аналитическом описании геометрических объектов многомерного пространства, можно считать актуальной научно-практической задачей.

Так или иначе, с задачами геометрического моделирования и оптимизации процессов и явлений сталкивались многие отечественные и зарубежные учёные в области прикладной и инженерной геометрии, которым относятся: В. Я. Аносов [1], В. М. Верещаг [2], Н. С. Гумен [3], С. В. Клименко [4], В. М. Комяк [5], А. В. Найдыш [2, 6], В. П. Радищев [7], О. В. Сергейчук [8] и др. Каждый решал поставленные перед ним прикладные и инженерные задачи своим способом в силу доступного математического аппарата и технического обеспечения, что породило множество способов моделирования. Однако наиболее близкой, даже родственной работой, являются исследования Н. Д. Вертинской [9, 10]. В этих работах задача моделирования многофакторных процессов и явлений представляется в следующем виде: в результате экспериментальных исследований или статистических данных имеется дискретные значения зависимых или независимых друг от друга факторов и функции отклика. Требуется определить аналитическую зависимость функции отклика от влияющих на неё факторов. Геометрическая интерпретация этой задачи заключается в том, что в пространстве размерности задан набор фиксированных точек, через которые необходимо построить геометрический объект и найти его уравнение. Основным отличием предложенного метода от представленных выше работ является метод конструирования геометрических объектов, проходящих через наперёд заданные точки. Другим отличием является использование последова-

тельности точечных уравнений для аналитического описания геометрического объекта в БН-исчислении (точечное исчисление Балюбы-Найдыша [11, 12, 13]).

Обоснование выбора БН-исчисления как аппарата геометрического моделирования и оптимизации многофакторных процессов и явлений

Для аналитического описания геометрических объектов, проходящих через наперёд заданные точки, используется математический аппарат БН-исчисления, основанный на инвариантах параллельного проецирования. Основным элементом БН-исчисления является точка, которая характеризуется несколькими параметрами. Количество параметров, которые определяют точку в пространстве, зависит от размерности этого пространства. Любой геометрический объект является организованным множеством точек. Поэтому точечные уравнения, справедливы для пространства любой размерности. Эта особенность БН-исчисления даёт возможность представлять геометрические объекты в многомерном пространстве посредством совокупности проекций на координатные оси. Можно выделить следующие основные преимущества БН-исчисления, как аппарата геометрического моделирования и оптимизации многофакторных процессов и явлений:

1. Точечные уравнения геометрических объектов инвариантны относительно размерности пространства глобальной системы координат. То есть в качестве параметров выбираются такие, которые являются инвариантными относительно параллельного проецирования (например, простое отношение трех точек прямой). Поэтому точечные уравнения справедливы для пространства любой размерности.
2. БН-исчисление позволяет работать в локальном симплексе, а результат получать в глобальном симплексе. Причем, переход от локального симплекса к гло-

бальному осуществляется автоматически.

3. В БН-исчислении каждой геометрической операции соответствует аналитическая операция. Таким образом, БН-исчисление позволяет представить любой геометрический алгоритм построения в аналитическом виде, или в виде точечного уравнения, или в виде вычислительного алгоритма, который, по сути, является упорядоченным множеством точечных уравнений. Поэтому в точечных уравнениях сохраняется наглядный геометрический смысл параметров, который известен из геометрического алгоритма построения. Эта особенность была потеряна для многих параметрических уравнений геометрических объектов, которые были получены аналитически, без использования геометрического алгоритма построения.
4. Точечные уравнения, по своей сути, является символьной записью. Переходя к глобальной декартовой системе координат, точечные уравнения заменяются на систему однотипных параметрических уравнений, количество которых зависит от размерности пространства глобальной системы координат. Эта операция получила название покоординатного расчёта. Точечные уравнения и вычислительные алгоритмы, полученные на их основе, можно считать оптимизированными для распараллеливания потоков и использования многоядерных микропроцессоров, поскольку вычисление каждой отдельной проекции на координатную ось можно выполнять параллельно. При этом вид уравнения остаётся без изменений, меняются только координаты точек.
5. В БН-исчислении был разработан специальный метод подвижного симплекса [16], который позволяет конструировать геометрические объекты любой сложности с наперед заданными свойствами. Использование метода подвижного симплекса позволяет установить зависимость между несколькими фак-

торами, что, в свою очередь, позволяет моделировать многофакторные процессы и явления, учитывая не только независимые факторы, но и факторы, которые зависят один от другого.

Общая концепция метода многомерной интерполяции

Исследуя методологию решения задачи геометрического моделирования и оптимизации многофакторных процессов и явлений с помощью многомерной интерполяции, можно выделить следующую концептуальную последовательность действий:

1. Анализ, сортировка и предварительная обработка исходных данных.
2. Составление геометрической схемы конструирования геометрического объекта, так называемого дерева геометрической модели, который соответствует исходной экспериментальной или статистической информации.
3. Выбор дуг интерполирующих кривых, проходящих через наперед заданные точки, которые на основе дерева геометрической модели будут формировать геометрический объект.
4. Моделирование геометрического объекта в соответствии с деревом геометрической модели на основе дуг кривых, проходящих через наперед заданные точки, которые бывают 2 видов: фиксированные, являющиеся графическим отображением исходной экспериментальной или статистической информации, и текущие, которые своим движением заполняют пространство и тем самым формируют геометрический объект.
5. Проверка достоверности полученной геометрической модели.
6. Аппроксимация полученной модели более простым геометрическим объектом. Этот этап является не обязательным, но в некоторых случаях позволяет значительно сократить итоговые вычисления, оказывая при этом незначительное вли-

яние на достоверность полученной модели.

7. Определение экстремумов на поверхности геометрического объекта функции отклика, которые могут иметь особое значение для исследуемого процесса или явления.

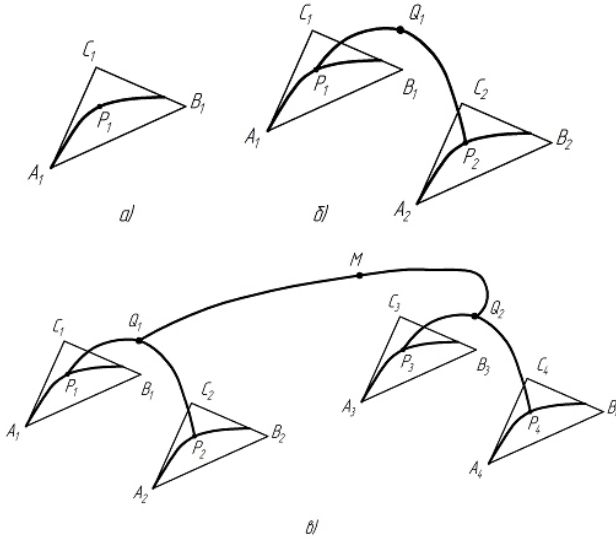


Рис. 1: Геометрическая сущность одно-, двух- и трёхфазного процесса

Кроме того можно отметить зависимость сложности геометрического объекта и размерности пространства, в котором он расположен, от количества моделируемых факторов. Так однофакторный процесс или явление можно представить однопараметрическим множеством точек — линией двумерного пространства, проходящей через наперед заданные точки, (рис. 1 а):

$$P_1 = (A_1 - C_1)p_1(u) + (B_1 - C_1)q_1(u) + C_1. \quad (1)$$

Точечное уравнение (1) представляет собой символьную запись. Выполнив по координатный расчёт, получим систему параметрических уравнений:

$$\begin{cases} x_{P_1} = (x_{A_1} - x_{C_1})p_1(u) + \\ \quad + (x_{B_1} - x_{C_1})q_1(u) + x_{C_1}, \\ y_{P_1} = (y_{A_1} - y_{C_1})p_1(u) + \\ \quad + (y_{B_1} - y_{C_1})q_1(u) + y_{C_1}. \end{cases}$$

Эта система представляет собой аналитическое описание проекций дуги плоской кри-

вой на оси глобальной системы координат. При этом оси $0x$ соответствует фактор, влияющий на процесс или явление, а оси $0y$ искомая функция отклика.

Двухфакторный — двухпараметрическим множеством точек — поверхностью трёхмерного пространства, проходящей через наперед заданные точки, (рис. 1 б):

$$\begin{cases} P_1 = (A_1 - C_1)p_1(u) + (B_1 - C_1)q_1(u) + C_1, \\ P_2 = (A_2 - C_2)p_2(u) + (B_2 - C_2)q_2(u) + C_2, \\ Q_1 = (P_1 - C_2)\varphi(\nu) + (P_2 - C_2)f(\nu) + C_2. \end{cases} \quad (2)$$

Трёхфакторный — трёхпараметрическим множеством точек — гиперповерхностью четырёхмерного пространства, проходящей через наперед заданные точки, (рис. 1 в):

$$\begin{cases} P_1 = (A_1 - C_1)p_1(u) + (B_1 - C_1)q_1(u) + C_1, \\ P_2 = (A_2 - C_2)p_2(u) + (B_2 - C_2)q_2(u) + C_2, \\ P_3 = (A_3 - C_3)p_3(u) + (B_3 - C_3)q_3(u) + C_3, \\ P_4 = (A_4 - C_4)p_4(u) + (B_4 - C_4)q_4(u) + C_4, \\ Q_1 = (P_1 - C_2)\varphi_1(\nu) + (P_2 - C_2)f_1(\nu) + C_2, \\ Q_2 = (P_3 - C_4)\varphi_2(\nu) + (P_4 - C_4)f_2(\nu) + C_4, \\ M = (Q_1 - C_4)\mu(w) + (Q_2 - C_4)\eta(w) + C_4. \end{cases} \quad (3)$$

Обобщая можно смоделировать любой геометрический объект, который будет носителем исходных точек и необходимого количества параметров. Тем не менее, для обеспечения прохождения через наперёд заданные точки искомого геометрического объекта, нужно обеспечить прохождение через наперёд заданные точки линий, из которых состоит искомый геометрический объект. А любая линия в БН-исчислении — это организованное определённым образом однопараметрическое множество точек.

Конструирование дуг кривых проходящих через наперед заданные точки

Для аналитического описания геометрических объектов многомерного пространства необходимо иметь специальный набор уравнений дуг кривых, проходящих через наперёд заданные точки. В данном случае возможно два принципиальных подхода: использование единой кривой и составной

кривой. Составные кривые получили широкое распространение в инженерной практике и различных системах автоматизированного проектирования в виде различного рода сплайнов и выпуклых обводов различного порядка гладкости. Однако для аналитического описания геометрических объектов многомерного пространства использование составных кривых является затруднительным и не всегда возможным. Это связано с тем, что моделирование составных геометрических объектов многомерного пространства возможно только на регулярной сети точек. К тому же сложность визуального восприятия многомерного пространства приводит к необходимости использования не зрительной, но логической наглядности, основанной на методах обобщения и аналогии. Использование единых дуг кривых, проходящих через наперед заданные точки, которые лишены приведенных выше недостатков, является более удобным и универсальным инструментом, позволяющим унифицировать полученные геометрические модели.

В работе [14] был предложен способ определения дуг кривых n -го порядка, которые проходят через $n + 1$ точек, на основе полиномов Бернштейна, который получил дальнейшее развитие в работе [15]. Рассмотрим более подробно сущность предложенного способа. Пусть задана ломаная линия $A_1A_2 \dots A_n, A_{n+1}$. Тогда дуга кривой n -го порядка имеет следующее точечное уравнение:

$$M = A_1 \bar{u}^n + A_2 C_n^1 \bar{u}^{n-1} u + A_3 C_n^2 \bar{u}^{n-2} u^2 + \dots + A_n C_n^{n-1} \bar{u} u^{n-1} + A_{n+1} u^n = \sum_{k=0}^n A_{k+1} C_n^k u^k \bar{u}^{n-k}, \quad (4)$$

где: C_n^1, \dots, C_n^k ; — коэффициенты Бернштейна; $\bar{u} = 1 - u$.

Рассмотрим частный случай, для которого $u(t) = t$ и точки ломаной A_i принадлежат одной плоскости, тогда получим дугу кривой Бернштейна, которая описывается точечным уравнением (4). В общем случае

кривые вида $M = \sum_{k=0}^n A_{k+1} C_n^k \bar{t}^{n-k}$, где $\bar{t} = 1 - t$, которые находятся в пространстве размерности n , являются кривыми $(n - 1)$ -й кривизны и в БН-исчислении получили название парабол n -го порядка.

Переопределим точки ломаной линии $A_1A_2 \dots A_n, A_{n+1}$ через точки $M_1M_2 \dots M_n, M_{n+1}$, которые принадлежат дуге кривой, определенной с помощью уравнения (4) при $u(t) = t$, изменяя соответствующим образом значение параметра t от 0 до 1. При значении параметра $t = 0$ имеем начало дуги кривой — точку M_1 , а при $t = 1$ — конец дуги кривой — точку M_{n+1} . Особенностью такого определения является то, что значение параметра распределяется равномерно: $t_k = \frac{k}{n}$, где k меняется от 0 до n . Причем точке M_k соответствует значение параметра t_k . Таким образом, получим систему $n + 1$ уравнений с неизвестными $A_1A_2 \dots A_n, A_{n+1}$. Решив эту систему уравнений методом Крамера и подставив в исходное уравнение, получим уравнение дуги кривой, проходящей через наперед заданные точки $M_1M_2 \dots M_n, M_{n+1}$.

Используя предложенный способ, были получены:

1. Дуга кривой 2-го порядка, проходящая через 3 наперед заданные точки:

$$M = M_1 \bar{t}(1 - 2t) + 4 \bar{t} t M_2 + M_3 t(2t - 1).$$

2. Дуга кривой 3-го порядка, проходящая через 4 наперед заданные точки:

$$M = M_1(\bar{t}^3 - 2,5 \bar{t}^2 t + \bar{t} t^2) + M_2(9 \bar{t}^2 t - 4,5 \bar{t} t^2) + M_3(-4,5 \bar{t}^2 t + 9 \bar{t} t^2) + M_4(\bar{t}^2 t - 2,5 \bar{t} t^2 + t^3). \quad (5)$$

3. Дуга кривой 4-го порядка, проходящая через 5 наперед заданных точек:

$$\begin{aligned}
 M = & M_1 \left(\bar{t}^4 - \frac{13}{3} \bar{t}^3 t + \frac{13}{3} \bar{t}^2 t^2 - \bar{t} t^3 \right) + \\
 & + M_2 \left(16 \bar{t}^3 t - \frac{64}{3} \bar{t}^2 t^2 + \frac{16}{3} \bar{t} t^3 \right) + \\
 & + M_3 \left(-12 \bar{t}^3 t + 40 \bar{t}^2 t^2 - 12 \bar{t} t^3 \right) + \\
 & + M_4 \left(\frac{16}{3} \bar{t}^3 t - \frac{64}{3} \bar{t}^2 t^2 + 16 \bar{t} t^3 \right) + \\
 & + M_5 \left(\bar{t}^3 t + \frac{13}{3} \bar{t}^2 t^2 - \frac{13}{3} \bar{t} t^3 + t^4 \right).
 \end{aligned}$$

4. Дуга кривой 5-го порядка, проходящая через 6 наперед заданных точек:

$$\begin{aligned}
 M = & M_1 \left(\bar{t}^5 - 6, 4 \bar{t}^4 t + 11, 2 \bar{t}^3 t^2 - \right. \\
 & \left. - 6, 4 \bar{t}^2 t^3 + \bar{t} t^4 \right) + \\
 & + M_2 \left(25 \bar{t}^4 t - 60, 4 \bar{t}^3 t^2 + \right. \\
 & \left. + 38, 5 \bar{t}^2 t^3 - 6, 3 \bar{t} t^4 \right) + \\
 & + M_3 \left(-25 \bar{t}^4 t + 122, 9 \bar{t}^3 t^2 - \right. \\
 & \left. - 95, 8 \bar{t}^2 t^3 + 16, 7 \bar{t} t^4 \right) + \\
 & + M_4 \left(-16, 7 \bar{t}^4 t - 95, 8 \bar{t}^3 t^2 + \right. \\
 & \left. + 122, 9 \bar{t}^2 t^3 - 25 \bar{t} t^4 \right) + \\
 & + M_5 \left(-6, 3 \bar{t}^4 t + 38, 5 \bar{t}^3 t^2 - \right. \\
 & \left. - 60, 4 \bar{t}^2 t^3 + 25 \bar{t} t^4 \right) + \\
 & + M_6 \left(\bar{t}^4 t - 6, 4 \bar{t}^3 t^2 + \right. \\
 & \left. + 11, 2 \bar{t}^2 t^3 - 6, 4 \bar{t} t^4 + t^5 \right).
 \end{aligned}$$

5. Дуга кривой 6-го порядка, проходящая через 7 наперед заданных точек:

$$\begin{aligned}
 M = & M_1 \left(\bar{t}^6 - 8, 7 \bar{t}^5 t + 22, 7 \bar{t}^4 t^2 - \right. \\
 & \left. - 22, 7 \bar{t}^3 t^3 + 8, 7 \bar{t}^2 t^4 - \bar{t} t^5 \right) + \\
 & + M_2 \left(36 \bar{t}^5 t - 133, 2 \bar{t}^4 t^2 + \right. \\
 & \left. + 151, 2 \bar{t}^3 t^3 - 61, 2 \bar{t}^2 t^4 + 7, 2 \bar{t} t^5 \right) + \\
 & + M_3 \left(-45 \bar{t}^5 t + 301, 5 \bar{t}^4 t^2 - \right. \\
 & \left. - 418, 5 \bar{t}^3 t^3 + 184, 5 \bar{t}^2 t^4 - 22, 5 \bar{t} t^5 \right) + \\
 & + M_4 \left(40 \bar{t}^5 t - 308 \bar{t}^4 t^2 + 600 \bar{t}^3 t^3 - \right. \\
 & \left. - 308 \bar{t}^2 t^4 + 40 \bar{t} t^5 \right) + \\
 & + M_5 \left(-22, 5 \bar{t}^5 t + 184, 5 \bar{t}^4 t^2 - \right. \\
 & \left. - 418, 5 \bar{t}^3 t^3 + 301, 5 \bar{t}^2 t^4 - 45 \bar{t} t^5 \right) + \\
 & + M_6 \left(7, 2 \bar{t}^5 t - 61, 2 \bar{t}^4 t^2 + \right. \\
 & \left. + 151, 2 \bar{t}^3 t^3 - 133, 2 \bar{t}^2 t^4 + 36 \bar{t} t^5 \right) + \\
 & + M_7 \left(-\bar{t}^5 t + 8, 7 \bar{t}^4 t^2 - 22, 7 \bar{t}^3 t^3 + \right. \\
 & \left. + 22, 7 \bar{t}^2 t^4 - 8, 7 \bar{t} t^5 + t^6 \right).
 \end{aligned}$$

Предложенные точечные уравнения являются универсальным инструментом и могут быть использованы в пространстве любой размерности. Но это не единственный возможный способ определения дуг кривых проходящих через наперед заданные точки. Например, в работе [14] приводится геометрический алгоритм, основанный на проективных свойствах кривой 2-го порядка, который позволяет определять дуги кривых 2-го порядка, проходящие через 3 наперед заданные точки и имеющие две касательные в начальной и конечной точках:

$$\begin{aligned}
 M = & (A - K) \frac{k \bar{t}(1 - 2t)}{k(1 - 2t)^2 + 2t \bar{t}} + \\
 & + (B - K) \frac{kt(2t - 1)}{k(1 - 2t)^2 + 2t \bar{t}} + K.
 \end{aligned}$$

Изменяя в этом уравнении параметр k , который представляет собой отношение на медиане, можно моделировать процессы и явления с помощью параболической, эллиптической и гиперболической интерполяции.

Кроме того в работах [15, 16] приводятся различные модифицированные геометрические алгоритмы, основанные на теоремах

Паскаля и Брианшона, и точечные уравнения дуг кривых 2-го порядка проходящих через 5 наперед заданных точек.

Заключение

В заключении хотелось бы отметить широкий спектр инженерных и прикладных задач, решение которых возможно с помощью моделирования многофакторных процессов и явлений методом многомерной интерполяции, используя библиотеку дуг кривых проходящих через наперед заданные точки. За последние 10 лет автором решены различные практические задачи по моделированию и оптимизации процессов и явлений в области строительных материалов и изделий [14, 17, 18, 19], теплотехники и кондиционирования [20], светотехники [21], экономики, пожарной и промышленной безопасности (некоторые из публикаций сданы в печать и не имеют пока ссылок на опубликованные труды). Кроме того частным случаем геометрического моделирования процессов и явлений является модификация метода конечных элементов, которая позволяет решать дифференциальные уравнения с частными производными с помощью многомерной криволинейной интерполяции.

Литература

- [1] Аносов В.Я. Геометрия химических диаграмм двойных систем [Текст] / В.Я. Аносов // М.: Изво АН СССР, 1959. — 188 с.
- [2] Принципи систематизації компонентів процесу для створення його геометричної моделі у БН-численні [Текст] / Є.В. Конопацький, В.М. Верещага, А.В. Найдиш, І.Г. Балюба // Науковий вісник Мелітопольського державного педагогічного університету імені Богдана Хмельницького. Серія: Математика. Геометрія. Інформатика. — Мелітополь: МДПУ ім. Б. Хмельницького, 2014. — Т.1. — С.75-79.
- [3] Гумен Н.С. Геометрические основы теории многообразий евклидоваго n -пространства применительно к геометрическому моделированию многопараметрических систем [Текст]: автореф. дис... доктора техн. наук: 05.01.01 / Гумен Н.С. — К., 1992. — 53 с.
- [4] Клименко С.В. Визуализация физических явлений и математических моделей в виртуальном окружении [Текст] / С.В. Клименко // отчет о НИР № 96-01-01273 (Российский фонд фундаментальных исследований). М.: ИСП РАН, 1998. — Точка доступа: https://elibrary.ru/download/elibrary_752361_58851754.htm.
- [5] Комяк В.М. Математическое и компьютерное моделирование нерегулярного размещения плоских геометрических объектов в областях произвольной пространственной формы [Текст]: автореф. дис. ... доктора техн. наук: 05.13.16 / Комяк В.М. — Харьков, 1996. — 43 с.
- [6] Найдыш А.В. Теоретические основы геометрического моделирования физико-механических свойств асфальтобетонов методами БН-исчисления [Текст] / Найдыш А.В., Конопацкий Е.В., Бумага А.И. // Науковий вісник Мелітопольського державного педагогічного університету імені Богдана Хмельницького. Серія: Математика. Геометрія. Інформатика. — Мелітополь: МДПУ ім. Б. Хмельницького, 2014. — Т.1. — С.111-117.
- [7] Радищев В.П. О применении геометрии четырех измерений к построению равновесных физико-химических диаграмм [Текст] / В.П. Радищев // Изв. Сектора физ.-хим. анализа. — 1947. — Т. 15. — С. 129-134.
- [8] Сергейчук О.В. Геометричне моделювання фізичних процесів при оптимізації форми енергоєфективних будинків [Текст]: автореф. дис. ... доктора техн. наук: 05.01.01 / Сергейчук О.В. — Київ, 2008. — 39 с.
- [9] Вертинская Н.Д. Теория нелинейных многомерных моноидальных поверхностей и её приложения [Текст]: автореф. дис. ... доктора техн. наук: 05.01.01 / Вертинская Н.Д. — Иркутск, 2006. — 31 с.
- [10] Вертинская Н.Д. Математическое моделирование многофакторных и многопараметрических процессов в многокомпонентных системах на базе конструктивной геометрии [Текст] / Н.Д. Вертинская // — Иркутск: ИрГТУ, 2009. — Часть 1 (Лекции) — 230 с.
- [11] Балюба, И. Г. Конструктивная геометрия многообразий в точечном исчислении [Текст]: дис. ... доктора техн. наук: 05.01.01 / Балюба И. Г. — Макеевка, 1995. — 227 с.
- [12] Найдыш В. М. Алгебра БН-исчисления [Текст] / В.М. Найдыш, И.Г. Балюба, В.М. Верещага // Прикладна геометрія та інженерна графіка: Міжвідомчий науково-технічний збірник. — К.: КНУБА, 2012. — Вип. 90. — С. 210-215.
- [13] Балюба И. Г. Точечное исчисление [Текст]: [учебное пособие] / И.Г. Балюба, В.М. Найдыш; под ред. В.М. Верещаги. — Мелітополь: МГПУ ім. Б. Хмельницького, 2015. — 236 с.
- [14] Конопацький, Є. В. Геометричне моделювання алгебраїчних кривих та їх використання при конструюванні поверхонь у точковому численні Балюби-Найдиша [Текст]: дис. ... канд. техн. на-

- ук: 05.01.01 / Є. В. Конопацький. — Мелітополь, 2012. — 164 с.
- [15] Бумага А. И. Геометрическое моделирование физико-механических свойств композиционных строительных материалов в БН-исчислении [Текст]: дис. ... канд. техн. наук: 05.23.05, 05.01.01 / Бумага А. И. — Макеевка, 2016. — 164 с.
- [16] Давыденко, И.П. Конструирование поверхностей пространственных форм методом подвижного симплекса: Дис. ... канд. техн. наук: 05.01.01. / И.П. Давыденко. — Макеевка, 2012. — 186 с.
- [17] Бумага А.И. Оптимизация состава комбинированного заполнителя мелкозернистого бетона методами БН-исчисления [Текст] / Бумага А.И., Братчун В.И., Конопацкий Е.В. // Современное промышленное и гражданское строительство. — Т.12, №2. — С. 92-98.
- [18] Конопацкий Е.В. Геометрическая модель зависимости предела прочности при сжатии модифицированного мелкозернистого дегтебетона от четырех параметров [Текст] / Конопацкий Е.В., Бумага А.И., Бочоришвили В.А. // Вестник Донбасской национальной академии строительства и архитектуры. Современные строительные материалы: сб. науч. тр. Макеевка: ДонНАСА, 2016. — Вып. 2016-1(117). — С. 55-61.
- [19] Конопацкий Е.В. Геометрическая модель процесса распределения прочностных характеристик в бетонной колонне [Текст] / Конопацкий Е.В., Воронова О.С. // Прикладная математика и вопросы управления. — Пермь: ПНИПУ, 2017. — № 1 — С.37-44.
- [20] Воронова О.С. Геометрическое моделирование физических параметров влажного воздуха [Текст] / Воронова О.С., Конопацкий Е.В. // Инвестиции, строительство, недвижимость как материальный базис модернизации и инновационного развития экономики: материалы VIII Международной научно-практической конференции, 13–15 марта 2018 г.: в 2 ч. — Томск: Изд-во Том. гос. архит.-строит. ун-та, 2018. — Ч.1. — С.318-322.
- [21] Егорченков, В.А. Принципы построения модели светового поля помещения с криволинейным четырёхугольным светопроемом с использованием точечного исчисления [Текст] / Егорченков В.А., Конопацкий Е.В. // Светотехника. — М.: ВНИСИ, 2015. — № 2. — С.59-60.

Формирование предварительной каркасной модели 3D-объекта по техническому чертежу, содержащему разрезы и сечения*

В. И. Дергунов, М. В. Лагунова, М. М. Смичёк, В. А. Тюрина

val.iv.dergunov@mail.ru | mvlenn@mail.ru | mariasmuchek@gmail.com | 55555_73@mail.ru

Нижегородский государственный архитектурно-строительный университет

Рассмотрена проблема автоматизированного синтеза предварительной каркасной модели трёхмерного объекта по техническому чертежу, содержащему разрезы и сечения. Описанный процесс позволяет расширить возможности алгоритма синтеза электронной модели 3D объекта по данным технического чертежа, разработываемого нижегородской научной геометрической школой под руководством профессора Роткова С.И. Сложности формализации анализа изображений, составляющих содержание технического чертежа, накладывают ограничения на входную информацию алгоритмов, решающих задачу синтеза, поэтому работа по расширению классов входной информации является актуальной и востребованной.

Ключевые слова: электронная геометрическая модель изделия, каркасная модель, предварительная каркасная модель, технический чертеж, разрезы и сечения, автоматизированное проектирование.

Formation of preliminary frame model of 3D-object according to the technical drawing, containing cuts and sections*

V. I. Dergunov, M. V. Lagunova, M. M. Smychek, V. A. Tyurina

Nizhny Novgorod State University of Architecture and Civil Engineering

The problem of automated synthesis of the preliminary wireframe model of a three-dimensional object is considered according to a technical drawing containing sections and sections. The described process makes it possible to expand the possibilities of the algorithm for synthesizing the electronic model of a 3D object according to the technical drawing developed by the Nizhny Novgorod scientific geometric school under the direction of Professor Rotkov S.I. The difficulties in formalizing the analysis of images that make up the contents of a technical drawing impose serious limitations on the input information of algorithms that solve the synthesis problem, therefore the work on expanding the classes of input information is very relevant.

Keywords: electronic geometric model of the product, wireframe model, preliminary wireframe model, technical drawing, sections and sections, computer-aided design.

1. Введение

Проблема синтеза 3D модели трёхмерного объекта по изображениям, составляющим многовидовый технический чертеж, является весьма актуальной в условиях развития различных информационных технологий, использующих 3D модель для решения различных научных, технических и технологических задач. Данная задача является объектом исследования нижегородской геометрической школы длительный период времени, начиная с работ академика В.С.Полозова и его учеников [1, 2, 3] и до настоящего времени, когда руководство этой научной проблематикой возглавляет профессор С.И. Ротков [4, 5].

Несмотря на бурное развитие современных технологий, сохраняются противоре-

чия между проблемой ввода геометрических данных для формирования электронной модели изделия и функциональными возможностями современных программных систем. Необходимо отметить факт, что на предприятиях различных отраслей промышленности накопились значительные объемы архивов чертежно-конструкторской документации на бумажных и электронных носителях, без преобразования данных с которых в 3D модели разных типов невозможно полноценное использование современных технологий.

Анализ литературных источников показывает, что обсуждаемая в статье проблема синтеза 3D модели до конца не решена ни в алгоритмическом, ни в практическом аспекте в виде устойчиво работающего промышленного программного продукта, и поэтому различные варианты решения этой трудно формализуемой обратной

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 17-07-00543, 18-07-20030.

задачи геометрии и графики являются актуальными и представляют теоретическую значимость [6].

Имеющиеся в различных системах функциональные операторы геометрического моделирования не покрывают всего многообразия ситуаций, возникающих в реальном проектировании, и вынуждают пользователя работать в ограниченных условиях, в том числе, ввода данных сразу в 3D пространстве, что противоречит методам и алгоритмам решения различных геометрических и графических задач.

Ранее разработанные научной нижегородской геометрической школой алгоритмы восстановления образа объекта использовались в качестве входной информации два или три стандартных вида объекта (вид спереди, вид сверху, вид слева), допускающие присутствие на них невидимых линий. При реальной работе с техническими чертежами конструкторы и проектировщики постоянно имеют дело с чертежами, включающими в свой состав разрезы и сечения, регламентируемые государственными стандартами ЕСКД [7, 8, 9]. Автоматизированный анализ таких изображений существенно расширяет возможности систем, позволяющих формировать электронную 3D модель изделия по техническому чертежу и пополнять банк моделей изделиями, ранее недоступными для восстановления.

2. Постановка задачи

Каркасная модель, автоматически построенная по информации, содержащейся на техническом чертеже, почти обязательно несет в себе ложную геометрическую информацию, исключения представляют лишь простейшие случаи [10]. Этот факт предопределен наличием на ортогональных проекциях объекта конкурирующих элементов. Поэтому каркасную модель, полученную автоматизированным путем, будем называть предварительной каркасной моделью (ПКМ). Для дальнейшей работы с ПКМ (преобразования ее в граничную модель, визуализации, проведения инженер-

ных расчетов и т.д.) необходимо уметь отделять ложную информацию от истинной. Поэтому математическое описание предварительной каркасной модели должно включать в себя не только определение графа (топология), но и всю полезную геометрическую информацию, которая может быть выявлена во время чтения чертежа.

Исходя из всего вышесказанного предлагается следующая структура ПКМ:

$$K = \{V, L, R\},$$

где:

$V = \{Vi = (Xi, Yi, Zi) \text{ из } R^3, AVi - \text{ атрибут, } i = 1 \dots Nv\}$ – множество вершин модели;

$L = \{Lj = (Xj(t), Yj(t), Zj(t)) \text{ из } R^3, Toj \leq t \leq Tj, ALj = \text{ атрибут, } j = 1 \dots NL\}$ – множество несущих линий, заданных параметрически;

$R = \{Rk = (V1k, V2k, Lk), V1k, V2k \text{ из } V, Lk \text{ из } L, ARk - \text{ атрибут, } k = 1 \dots NR\}$ – множество ребер модели.

Рассмотрим атрибуты этих множеств, так как именно они позволяют выявлять ложную геометрическую информацию.

Атрибут вершины $AVi = \{0, 1\}$, где 0 – внутренняя вершина каркаса, 1 – внешняя вершина.

Будем называть внешней вершиной каркаса такую вершину, у образа которой хотя бы на одной из проекций отсутствуют конкурирующие точки. Если же на каждом проекционном изображении у образа вершины есть конкурирующие точки, то такую вершину назовем внутренней вершиной каркаса. Внешние вершины являются истинными элементами каркаса, а внутренние вершины каркаса являются подозрительными на принадлежность к ложным элементам каркаса. Заметим, что и ребра, инцидентные внутренним вершинам каркаса, являются подозрительными на принадлежность к ложным элементам.

Атрибут линии $ALj = \{0, 1, 2, 3, 4, 5, 6\}$, где 0 – сплошная основная линия, 1 – сплош-

ная тонкая, 2 – штриховая, 3 – штрихпунктирная тонкая, 4 – разомкнутая, 5 – тонкая волнистая, 6 – сплошная тонкая с изломом.

Атрибут ребра $Ark = \{0, 1\}$, где 0 – конструктивное ребро, 1 – неконструктивное ребро.

Если образы вершин на каждом проекционном изображении принадлежат некоторому общему ребру и построенные для этого ребра проецирующие фигуры пересекаются по некоторой пространственной кривой, то участок этой кривой между данными вершинами будем называть конструктивным ребром. Если образы вершин смежны только на одном проекционном изображении, то будем считать, что между ними может быть построено неконструктивное ребро.

При построении каркасной модели указанные выше атрибуты легко получаемы, поэтому входную информацию для задачи преобразования каркаса можно требовать именно в таком виде.

Данную математическую модель каркаса легко формализовать на любом объектно-ориентированном языке программирования. Топологическая информация легко считывается и для удобства работы алгоритма может храниться в списковой структуре. Список, в котором хранится вся информация о графе имеет следующий вид:

$\{Vi - \text{вершина из } V, Ti - \text{атрибут истинности вершины}, Ki - \text{валентность вершины}, V1i, \dots, Vki - \text{набор смежных ей вершин из } V, \text{ где } k = 1 \dots Ki, i = 1 \dots NV\}$.

Атрибут истинности $i = \{0, 1\}$, где 0 – если вершина истинная, 1 – подозрительная на ложную.

Геометрическая информация также легко считывается из структуры каркасной модели. К ней приходится обращаться в двух случаях:

- если в ходе преобразования каркаса выяснилось, что он содержит ложные геометрические элементы и необходимо их выявить;

- если в ходе преобразования каркаса удалось сформировать набор граней объекта (т.е. ложных элементов нет), и необходимо выделить поверхности, ограничивающие искомый 3-мерный объект.

Таким образом, описанная математическая модель каркаса удобна для осуществления дальнейших преобразований каркасной модели объекта (например, в граничную модель).

В данной статье предлагается расширить спектр обрабатываемой входной информации, содержащейся на техническом чертеже, а именно, за счёт возможности анализировать информацию о разрезах и сечениях объекта. В результате проведенного исследования было выявлено, что формирование предварительной каркасной модели по чертежу, на котором присутствуют разрезы, дает возможность получать предварительную каркасную модель с большим количеством истинных элементов (вершин и ребер), и позволяет сократить количество процедур по распознаванию и удалению ложных геометрических элементов, присутствующих в ПКМ.

3. Общая блок-схема процесса

Рассматриваемая нами задача относится, по классификации В. С. Полозова, к неформальным задачам, т.е. таким, общий алгоритм решения которых неизвестен, но в каждом конкретном случае они могут быть разрешены человеком на основе опыта и интуиции [1]. Поэтому для того, чтобы перейти к алгоритмам, позволяющим реализовать возможность автоматического построения 3D модели по чертежу, содержащему разрезы и сечения, необходимо построить формальные модели процессов чтения такого чертежа человеком. Первым этапом построения формальной модели неформальной задачи является изучение поведения человека в ходе ее решения. Процесс описания поведения человека в ходе решения неформальной задачи принято называть эвристическим моделированием.

ем. Такой подход делает возможным выявить исходные данные задачи и наложенные на них ограничения, определить признаки правильности решения и т.д. Такой подход был применен авторами данной работы для составления схемы алгоритма [11] чтения чертежа, содержащего разрезы и сечения (результат представлен на рис. 1).

Проведенные наблюдения за поведением человека, перед которым стоит задача «прочитать» чертеж, позволили выделить следующие признаки, присутствующие на чертеже и позволяющие выявить изображения, на которых присутствуют разрезы и сечения:

1. Наличие на чертеже объекта «Штриховка».
2. Наличие на чертеже буквенно-цифровых обозначений типа «А-А» или «1-1».
3. Наличие на изображениях чертежа линий обрыва.
4. Наличие на чертеже условного графического обозначения «повернуто».
5. Наличие на чертеже изображения в разрыве между частями одного и того же вида.
6. Наличие на чертеже изображения, связанного с видом нормальной штрихпунктирной тонкой линией.

Для определения типов разрезов и сечений человек, осуществляющий «чтение» чертежа, в первую очередь обращает внимание на следующие признаки:

1. Способ оформления области штриховки.
2. Обозначения секущей плоскости.
3. Как расположен след секущей плоскости относительно системы координат.
4. Какова структура следа секущей плоскости.
5. Количество секущих плоскостей.

Анализ данных признаков позволяет определить виды разрезов и сечений, присутствующих на чертеже, понять их назначение и, в результате, получить полное представление об особенностях формы изображенного на чертеже объекта.

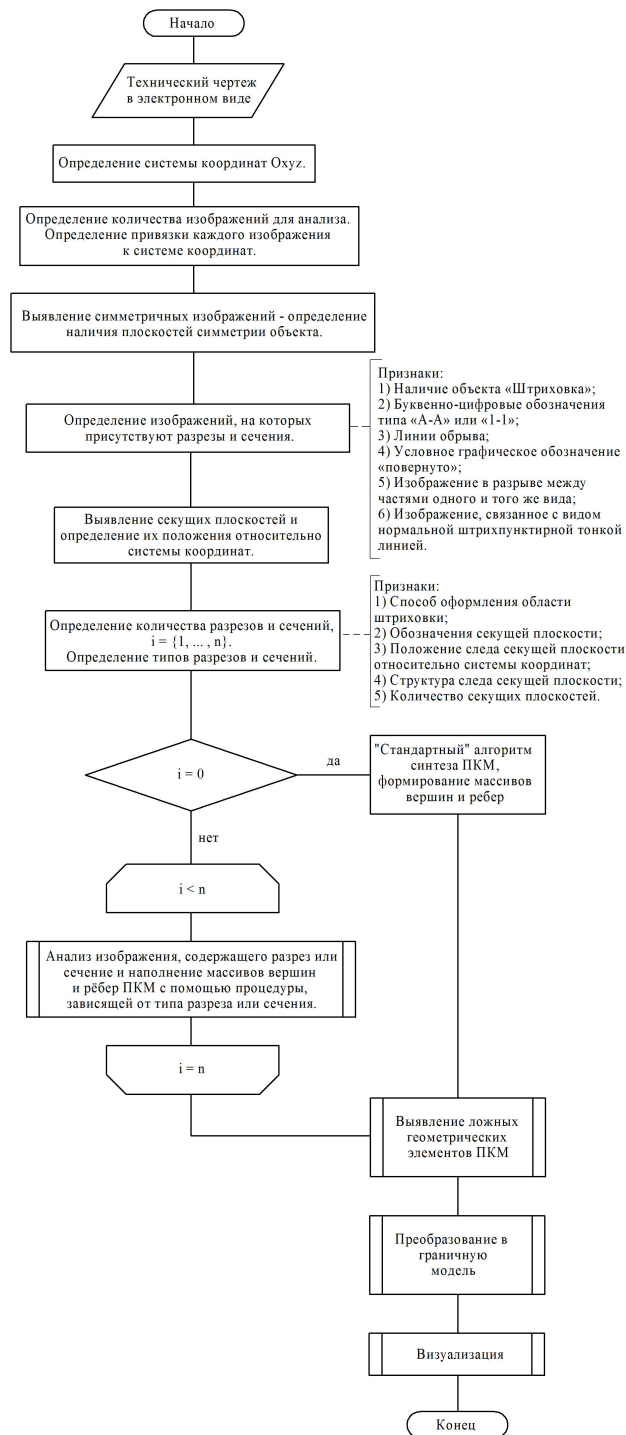


Рис. 1: Схема алгоритма чтения чертежа, содержащего разрезы и сечения. Большой интерес с точки зрения постановки эвристического эксперимента представляют из себя блоки «Определение изображений, на которых присутствуют разрезы и сечения» и «Определение типов разрезов и сечений»

Электронный чертеж объекта, выполненный в той или иной системе геометрического моделирования, позволяет получить информацию о приведенных выше признаках

в формальном (цифровом) представлении и дает возможность после процедур преобразования полученных данных перейти к построению предварительной каркасной модели объекта в автоматизированном режиме.

4. Заключение

Работа над программной реализацией описанного подхода к решению поставленной задачи требует подробного эвристического анализа каждого шага алгоритма и приведения его к наиболее полному формальному виду. Процедуры преобразования данных электронного чертежа объекта могут отличаться конкретной формой реализации, зависящей от вида системы, используемой для построения чертежа, но общая схема синтеза электронной 3D модели от вида системы не зависит и в перспективе может быть использована и для работы с растровыми архивными изображениями.

Об авторах

Дергунов Валентин Иванович, к.т.н., профессор, профессор кафедры Инженерной геометрии, компьютерной графики и автоматизированного проектирования Нижегородского государственного архитектурно-строительного университета.
val.iv.dergunov@mail.ru

Лагунова Марина Викторовна, д.п.н., профессор, профессор кафедры Инженерной геометрии, компьютерной графики и автоматизированного проектирования Нижегородского государственного архитектурно-строительного университета.
mvlmn@mail.ru

Смычѳк Мария Михайловна, ассистент кафедры Инженерной геометрии, компьютерной графики и автоматизированного проектирования Нижегородского государственного архитектурно-строительного университета.
mariasmychek@gmail.com

Тюрина Валерия Александровна, к.т.н., доцент, профессор кафедры Инженерной гео-

метрии, компьютерной графики и автоматизированного проектирования Нижегородского государственного архитектурно-строительного университета.
55555_73@mail.ru

Литература

- [1] Котов И.И. Алгоритмы машинной графики / И.И. Котов, В.С. Полозов, Л.В. Широкова. — М.: Машиностроение, 1977. — 232 с.
- [2] Полозов В.С. Моделирование и синтез операторов геометрического расчета и машинной графики в системах автоматизированного проектирования и автоматизации технологической подготовки производства: Автореф. дис. докт. тех. наук: 05.01.01. / В.С. Полозов. — М.: МАИ, 1983. — 41 с.
- [3] Полозов В.С. Геометрические и графические задачи: Автоматизированное проектирование / В.С. Полозов, О.А. Будеков, С.И. Ротков, Л.В. Широкова. — М.: Машиностроение, 1983. — 280 с.
- [4] Ротков С.И. Средства геометрического моделирования и компьютерной графики пространственных объектов для CALS-технологий: Дис. докт. техн. наук: 05.01.01 / С.И. Ротков. — Н. Новгород, 1999. — 287 с.
- [5] Зудин А.А. Моделирование пространственных объектов с криволинейными поверхностями методом чтения чертежа / А.А. Зудин, С.И. Ротков // Сб. докл. Всесоюзного научно-технического семинара «Машинная графика и автоматизация проектирования в радиоэлектронике» — Челябинск, 1990.
- [6] Мошкова Т.В. Проблема синтеза модели 3D объекта по его проекционным изображениям. Аналитический обзор / Т.В. Мошкова, С.И. Ротков, В.А. Тюрина // Научная визуализация, 2018, том 10, номер 1, стр. 135-156, DOI: 10.26583/sv.10.1.11
- [7] ЕСКД. Единая система конструкторской документации. ГОСТ 2.052-2015 Электронная модель изделия. Общие положения / База нормативных документов [Электронный ресурс] — Режим доступа: <http://russgost.ru>.
- [8] ЕСКД. Единая система конструкторской документации. ГОСТ 2.305-2008 Изображения: виды, разрезы, сечения. / База нормативных документов [Электронный ресурс] — Режим доступа: <http://russgost.ru>.
- [9] ЕСКД. Единая система конструкторской документации. ГОСТ 2.303-68 Линии. / База нормативных документов [Электронный ресурс] — Режим доступа: <http://russgost.ru>.
- [10] Тюрина В.А. Разработка методов преобразований каркасной модели в задаче синтеза образа 3D-объекта по его проекциям: Дисс. канд. техн. наук: 05.01.01 / В.А. Тюрина. — Н. Новгород, 2003. — 170 с.
- [11] ЕСПД. Единая система программной документации. ГОСТ 19.701-90 (ИСО 5807-85). Схемы алго-

ритмов, программ, данных и систем. Обозначения условные и правила выполнения. / База норма-

тивных документов [Электронный ресурс] — Режим доступа: <http://russgost.ru>.

Робастная векторизация графических документов с выраженной ориентацией линейных объектов*

Ю. Г. Васин¹, Д. Ю. Васин¹, В. П. Громов¹, С. И. Ротков²
dm04@list.ru | gvp1941@bk.ru | rotkov@nngasu.ru

¹Нижегородский Государственный Университет им. Н. И. Лобачевского

²Нижегородский Государственный Архитектурно-Строительный Университет

Рассмотрены вопросы развития комбинаторно-геометрического подхода в задачах векторизации большеформатных, сложноструктурированных пространственно-распределенных данных.

Ключевые слова: *большеформатные, сложноструктурированные пространственно-распределенные данные, комбинаторно-геометрический подход в задачах обработки видеоданных, иерархические структурные растровые и векторные модели, векторизация, растровые изображения, векторные изображения.*

Technology of creation of graphic databases of floor-by-floor plans of buildings*

Yu. G. Vasin¹, D. Yu. Vasin¹, V. P. Gromov¹, E. A. Kustov¹,
S. I. Rotkov², Yu. V. Yasakov¹

¹Nizhny Novgorod State University. N. I. Lobachevsky

²Nizhny Novgorod State Architectural and Construction University

Questions of development of combinatory and geometrical approach in problems of vectorization of the bolsheformatny, slozhnostrukturirovanny spatial distributed data are considered.

Keywords: *the bolsheformatny, slozhnostrukturirovanny spatial distributed data, combinatory and geometrical approach in problems of processing of video data, hierarchical structural raster and vector models, vectorization, raster images, vector images.*

Введение

В настоящее время актуальной остается проблема создания цифровых баз данных большеформатных, сложноструктурированных графических документов (БФГД). К этому классу отнесем планы и схемы поэтажной застройки зданий, конструкторскую документацию, технические чертежи, электрические схемы, топографические и морские навигационные карты и др. При этом источником, как правило, являются графические изображения на бумажных носителях, а создание цифровых документов необходимо обеспечить в терминах соответствующей проблемной области. Однако, значительный класс БФГД выполняется ручным способом со слабо формализованным начертанием объектов, произвольной ориентацией, произвольными размерами символов, наличием большого количества межобъектных наложений, примыка-

ний и пересечений, образуя подкласс БФГД со слабо формализованным описанием объектов (СФГД).

В процессе первичного анализа документов, относящихся к СФГД, выявляются характерные особенности, присущие данному классу документов, а именно: СФГД типа машиностроительных чертежей, схем, планов поэтажной застройки зданий являются документами массового типа; как правило, данный класс документов исполнен на бумажных носителях среднего и плохого качества; на планах и схемах отсутствует жесткий стандарт и часто не соблюдены правила отрисовки отдельных графических знаков; в зависимости от качества носителя СФГД и правильности выбора параметров оцифровки, степень зашумленности растрового изображения графического документа (РИГД) может варьироваться в широких пределах; метрическая точность графических знаков также варьируется в широком диапазоне значений (рис. 1).

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 17-07-00543, 18-07-00715, 18-07-20030.

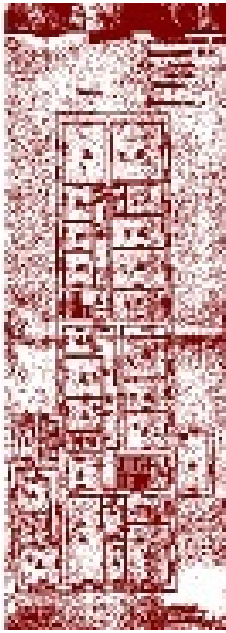


Рис. 1: Пример исходного графического документа типа плана поэтажной застройки здания

Низкоуровневая структурированная модель растровых данных

Пусть РИГД представлено в бинарном пиксельном формате, при этом функция яркости пикселей имеет вид:

$$f(x, y) = \begin{cases} 225, & \text{если пиксель фоновый;} \\ 0, & \text{если пиксель не фоновый.} \end{cases}$$

Непроизводным элементом модели является *штрих* – одномерный кластер связанных пикселей заданного цвета, представленный набором чисел $S = (t, N, K, C)$, где: t – номер строки растра; N, K – координаты начала и конца связанной одноцветной пиксельной области вдоль линии растра; C – код цвета связанных пикселей, образующих штрих. Для бинарных растровых изображений (БРИ) штрих описывается тройкой типа $S = (t, N, K)$, где: t – номер строки растра; N, K – координаты начала и конца штриха, определяемые соотношениями:

$$N = x \wedge f(x - 1, y) \wedge \bar{f}(x, y), \\ K = x \wedge \bar{f}(x, y) \wedge f(x + 1, y).$$

Совокупность штрихов для заданной строки растра t определяет строку штрихов $SH = \{S_i\} (i = 1, 2, \dots, m)$, где S_i – штрих.

Совокупность строк штрихов РИГД определяет его штриховую модель описания.

Топологические свойства штриха $S = (t, N, K)$ определяются через его связность со штрихами смежных строк. Два штриха S_t и S_{t-1} смежных строк t и $t - 1$ связаны, если $(N_{t-1} \geq N_t) \wedge (N_{t-1} \leq K_t) \vee (K_{t-1} \geq N_t) \wedge (K_{t-1} \leq K_t)$. Связный штрих – это набор $H = \{S, sw_p, sw_s\}$, sw_p, sw_s – количество связанных штрихов со штрихами смежных строк, а t – номер текущей строки штрихов растра. Тогда $sw_p = 0$, если в предыдущей строке $p = t - 1$ нет ни одного штриха, для которого выполняется условие $(N_p \geq N_t) \wedge (N_p \leq K_t) \vee (K_p \geq N_t) \wedge (K_p \leq K_t)$, иначе, $sw_p = \mu_p$, μ_p – кратность выполнения данного условия, то есть количество связанных штрихов предыдущей строки. Аналогично, $sw_s = 0$, если в следующей строке $s = t + 1$ нет ни одного штриха, для которого выполняется условие: $N_s \geq N_t \wedge N_s \leq K_t \vee K_s \geq N_t \wedge K_s \leq K_t$, иначе, $sw_s = \mu_s$, μ_s – кратность выполнения данного условия, то есть количество связанных штрихов последующей строки. Здесь (N, K) – координаты начала и конца исследуемого на связность штриха, (N_t, K_t) – координаты начала и конца текущего штриха предыдущей смежной строки штрихов, (N_s, K_s) – координаты начала и конца текущего штриха следующей смежной строки штрихов. По суперпозиции значений sw_p и sw_s можно ввести следующую классификацию графических ситуаций.

В текущей строке:

- штрих изолированный (Ши):
 $sw_p = 0 \wedge sw_s = 0$;
- штрих начала растрового объекта (Шн):
 $sw_p = 0 \wedge sw_s = 1$;
- штрих конца растрового объекта (Шк):
 $sw_p = 1 \wedge sw_s = 0$;
- штрих слияния растровых объектов (Шс):
 $sw_p > 1$;
- штрих расщепления растровых объектов (Шр):
 $sw_s > 1$;

- штрих слияния и расщепления растровых объектов (Шср):
 $sw_p > 1 \wedge sw_s > 1$.

Растровый простой объект (РПО) — кластер связанных штрихов, не содержащий графических ситуаций типа слияния и расщепления, то есть для любого штриха кластера выполняется условие $sw_p < 2 \wedge sw_s < 2$.

Растровый составной объект (РСО) — кластер связанных штрихов, для каждого штриха которого выполняется условие: $sw_p > 0 \vee sw_s > 0$, и для любых двух элементов (штрихов или пикселей) найдется хотя бы одна соединяющая их пиксельная траектория, состоящая из связанных смежных элементов.

Линейно-площадная модель РИГД — это представление исходного РИГД как логической суммы двух РИГД, один из которых состоит из пикселей *растровых линейных* объектов (РЛО), а второй — из пикселей *растровых площадных* объектов (РПО). Такое расщепление исходного РИГД на две компоненты возможно при задании каждому пикселю исходного РИГД свойства протяженной связности L_{sw} . Определим L_{sw} для пикселя как длину минимального штриха, полученного по всем возможным направлениям линии сканирования, проходящей через этот пиксель, то есть $L_{sw} = \min_u(L_u)$, $u = 1, 2, \dots, k$, где u — направление линии сканирования, L_u — длина штриха, которому принадлежит исследуемый пиксель. Сравнивая полученную величину с пороговой пиксельной толщиной линейных растровых объектов, можно осуществить бинарную классификацию на два класса: пиксель, принадлежащий РЛО, или пиксель, принадлежащий РПО, простым пороговым решающим правилом: если $L_{sw} < Por$, то пиксель принадлежит РЛО, иначе, пиксель принадлежит РПО.

Постановка задачи

Для СФГД типа чертежей, схем и планов характерно изображение сюжетов рисунка, с помощью отрезков линий с чет-

ко выраженным направлением их ориентации. Если при сканировании этих документов направление сканиции выбрать в соответствии с преимущественной ориентацией этих отрезков, то штриховая модель отсканированных данных позволяет построить простые и надежные алгоритмы векторизации и осуществить эффективное геометрическое моделирование растрового описания этих документов.

Векторизация является базовой операцией в большинстве систем обработки и анализа графической информации. Если растровое описание векторизуемых объектов не содержит искажений и шумов то существующие локальные алгоритмы векторизации вполне удовлетворительно справляются с этой задачей, хотя надо иметь ввиду, что получаемые при этом векторизованные объекты требуют дополнительного сглаживания либо аппроксимации, а это не всегда удается сделать успешно, то есть одновременно удовлетворить требованиям метрической точности аппроксимации и геометрической достоверности векторизованного объекта.

В работах [1, 2] предложен комбинаторно — геометрический подход в задачах обработки сложной графической информации. В рамках развития этого подхода в настоящей работе рассматриваются методы векторизации растровых графических изображений при наличии шумов в исходных данных и предложена геометрическая модель описания графических документов с выраженной ориентацией линейных объектов.

Алгоритмы формирования линейной модели РИГД

До начала векторизации предлагается выполнить этап предварительной обработки растровых данных, включающий следующие операции над исходным растром:

1. Определение преимущественно выраженного направления ориентации ЛРО, как направление наиболее протяженных ЛРО на данном РИГД.

2. Поворот исходного РИГД, таким образом, чтобы новое направление сканирования графического документа совпало или было близко к найденной на шаге 1 ориентации ЛРО.
3. Конвертация исходного РИГД в штриховую модель.
4. Разбиение полученного множества штрихов на набор РСО.
5. Формирование линейно-площадной модели РИГД в виде двух РИГД, содержащих наборы ЛРО и РПО исходного РИГД. Для этого исходный набор РСО сегментируем на кластеры связных коротких и длинных штрихов по пороговому значению $L_p = k \times d_{Line}$, где: d_{Line} – средняя толщина растровой линии, $k = \{1 - \text{горизонтальные}, 2 - \text{наклонные}, 3 - \text{вертикальные}\}$ – коэффициент наклона растровых линий на документе. Далее выполняем слияние кластеров заданного типа методом ε – кластеризации, то есть формируем РСО с заданными свойствами и полученные РСО считаем исходными растровыми объектами для тех или иных алгоритмов векторизации.

Как было отмечено, для РИГД типа поэтажных планов, схем, технических чертежей характерно наличие большого количества ЛРО. Поэтому далее остановимся подробно на алгоритмах векторизации именно этого вида растровых объектов. В работе исследовались два алгоритма векторизации ЛРО – глобальный и локальный.

В глобальном алгоритме очередной исходный ЛРО рассматривается как некоторое множество точек плоскости $P = \{x_i, y_i\}$, ($i = 1, 2, \dots, N$), где каждая точка $\{x_i, y_i\}$ является векторизованным описанием текущего пикселя обрабатываемого ЛРО. Для множества P методом наименьших квадратов (МНК) строится кусочно – линейная аппроксимирующая с заданным ε функция, которая и является выходным результатом векторизации текущего ЛРО.

В локальном алгоритме производится последовательное отслеживание узлов линии методом самоцентрируемого скользящего окна с использованием следующих алгоритмических операций:

1. Центр растрового окна заданного размера $H_w = d_{Line}$ помещается в текущий свободный пиксель ЛРО.
2. Для координат n пикселей, попавших в окно, вычисляется среднее значение и их линейная ориентация $\tilde{x} = \frac{1}{n} \sum_{i=1}^n x_i$; $\tilde{y} = \frac{1}{n} \sum_{i=1}^n y_i$; U_x, U_y , где: U_x, U_y – компоненты первого собственного вектора, полученного по МНК, и линейно аппроксимирующего множество пикселей окна.
3. Среднее значение принимаем за текущий узел векторизуемого ЛРО.
4. Перемещаем окно по вектору линейной ориентации до появления первого свободного пикселя.
5. Шаги 2 – 4 выполняем до тех пор, пока в перемещаемом окне будет хотя бы один свободный пиксель, в противном случае считаем, векторизацию текущего ЛРО законченной.

Для множества отрезков, полученных из текущего РСО, выполняется операция метрического и геометрического согласования метрики отрезков.

По согласованной метрике отрезков формируется и тоговая геометрическая модель (рис. 2).

Благодарности

Работа выполнена при финансовой поддержке Российского Фонда Фундаментальных Исследований, гранты РФФИ № 17-07-00543, 18-07-00715.

Литература

- [1] Васин Ю.Г., Башкиров О.А., Чудинович Б.М. Комбинаторно-геометрический подход в задачах анализа сложной графической информации. – Автоматизация обработки сложной графической информации: Межвуз. сб. науч. тр. / Под ред. Ю.Г. Васина: Горьк. гос. ун-т. Горький, 1987.

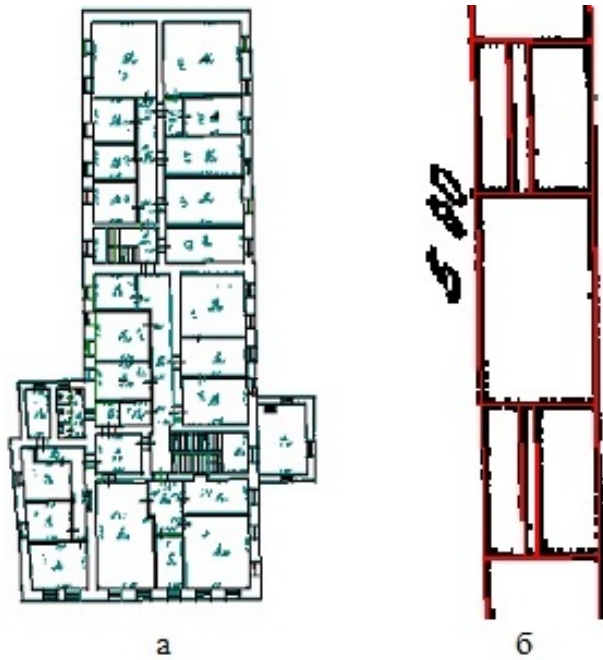


Рис. 2: а) итоговая геометрическая модель; б) ее увеличенный фрагмент. На базе изложенных алгоритмов создано оригинальное программное обеспечение, используемое при автоматическом цифровании схем поэтажной планировки в задачах создания электронных кадастров

- [2] Васин Ю.Г., Башкиров О.А., Рудометова С.Б. Математические модели структурированного описания графических изображений. В сб. Автоматизация обработки сложной графической информации, ГГУ, Горький, 1984, с. 92-110.

Технология создания графических баз данных поэтажных планов зданий*

Ю. Г. Васин¹, Д. Ю. Васин¹, В. П. Громов¹, Е. А. Кустов¹,
С. И. Ротков², Ю. В. Ясаков¹, С. В. Клименко²

dm04@list.ru | gvp1941@bk.ru | rotkov@nngasu.ru | kustov@pmk.unn.ru | yuri961@yandex.ru

¹Нижегородский Государственный Университет им. Н. И. Лобачевского

²Нижегородский Государственный Архитектурно-Строительный Университет

Рассмотрена автоматизированная технология создания графических баз данных поэтажных планов зданий на основе оригинального информационного обеспечения, моделей и структур данных, в рамках дальнейшего развития комбинаторно-геометрического подхода в задачах обработки большеформатных, сложноструктурированных пространственно-распределенных данных.

Ключевые слова: автоматизированная технология создания графических баз данных, автоматизация обработки поэтажных планов зданий, CALS-технологии, большеформатные, сложноструктурированные пространственно-распределенные данные.

Technology of creation of graphic databases of floor-by-floor plans of buildings*

Yu. G. Vasin¹, D. Yu. Vasin¹, V. P. Gromov¹, E. A. Kustov¹,
S. I. Rotkov², Yu. V. Yasakov¹, S. V. Klimenko²

¹Nizhny Novgorod State University. N. I. Lobachevsky

²Nizhny Novgorod State Architectural and Construction University

The automated technology of creation of graphic databases of floor-by-floor plans of buildings on the basis of original information support, models and structures of data, within further development of combinatorial and geometrical approach in problems of processing of the bolsheformatny, slozhnostrukturirovanny spatial distributed data is considered.

Keywords: the automated technology of creation of graphic databases, automation of processing of floor-by-floor plans of buildings, CALS technologies, the bolsheformatny, slozhnostrukturirovanny spatial distributed data.

Переход к цифровой экономике

Переход к цифровой экономике, реализация концепций PLM, ERP, умного дома, механизмов импортозамещения, диктуют расширение и дальнейшее внедрение информационных технологий в области строительства и архитектуры. На сегодняшний день сфера применения информационных технологий в строительной отрасли очень обширна — от создания простых чертежей и документов до компьютерного проектирования и управления строительным объектом. Проектирование зданий и сооружений все больше превращается в единый комплекс работ, включающий использование CALS — технологий на всех этапах жизненного цикла изделий, начиная от стадии разработки и заканчивая вводом в эксплуатацию гото-

вого объекта.

Необходимость внедрения информационной технологии для развития строительства объясняется требованиями к сокращению сроков проектирования и подготовки производства для выпуска новых и модернизируемых изделий, затрат на проектирование и производство, стоимости долговременного послепродажного обслуживания. Представленная технология является расширением разработанных ранее в НИИ ПМК ННГУ (ныне ИТММ ННГУ) технологией АРМ РАСТР — 2 и АРМ РАСТР — ТП и базируется на едином с ними информационно — терминологическом обеспечении, включающем:

- базу знаний;
- классификатор;
- модели и правила ввода метрической и семантической информации об объектах.

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 17-07-00543, 18-07-00715, 18-07-20030.

Система, обеспечивающая работоспособность технологии, представляет собой открытую систему, состоящую из отдельных модулей, связанных общей структурой данных, состав модулей системы можно дополнять и изменять.

Исходными документами для создания баз данных являются планы всех этажей здания на бумажных носителях. Анализ растровых изображений планов показал, что для них характерны следующие особенности:

- графические документы поэтажных планов относятся к документам массового типа и используются большинством хозяйственных служб;
- планы, как правило, выполнены на бумажных носителях;
- отсутствует жесткий стандарт либо часто не соблюдаются правила отрисовки отдельных графических знаков;
- базовым графическим элементом для изображения массовых знаков типа стен, дверных и оконных проемов, дверей, окон, границ помещений являются сегменты линейных объектов (линейные отрезки), образующие вкуче для одного документа совокупность параллельных и ортогональных линий;
- в зависимости от качества носителя графического документа и правильности выбора параметров сканирования исходных документов степень зашумленности растрового изображения может быть очень значительной;
- метрическая точность графических знаков может меняться в большом диапазоне значений;
- ярко выраженная топологическая нагрузка на базовые графические элементы линейные отрезки, то есть использование одного и того же линейного отрезка в описании нескольких графических знаков.

Процесс создания состоит из следующих технологических этапов:

1. Для всего комплекта поэтажных планов здания:

- оценка качества исходных документов;
 - выбор метода создания (полностью ручной ввод или с использованием процедур автоматического распознавания условных знаков). Выбор определяется нагрузкой документа и качеством прорисовки условных знаков;
 - в зависимости от выбранного метода создания графической базы данных осуществляется выбор параметров сканирования исходных документов;
 - сканирование планов всех этажей здания с получением их растрового представления в штриховом формате;
 - координатная привязка растровых изображений введенных документов и её контроль.
2. Дальнейшая обработка ведется последовательно для плана каждого этажа. Для обработки в первую очередь выбирается план того этажа, который содержит максимальное количество условных знаков, присутствующих на тех же местах на планах других этажей. Затем выполняются автоматические процедуры распознавания условных знаков и формирования соответствующих объектов в базе данных, включающие следующие технологические операции:
 - автоматизированное редактирование растровых изображений;
 - построение контурной модели изображения;
 - построение линейной модели изображения как совокупности горизонтальных и вертикальных растровых линий;
 - формирование сегментно-узловой модели изображения;
 - формирование сегментно-контурной модели изображения;
 - кластеризация простых контуров, разрезание сложных контуров, распознавание графических знаков: «помещение», «элемент стены», «двер-

- ной проем», «оконный проем», «дверь», «окно», «стенной шкаф», «лестница» и лестничный пролет; топологическое согласование метрики распознанных графических знаков; пересчет объектов в систему координат создаваемой базы данных;
- загрузка в базу данных семантики и метрики распознанных объектов;
 - интерактивный просмотр и редактирование полученной базы данных и, в случае необходимости, ее дополнение нераспознанными объектами по моделям.
3. Очередной технологический этап связан с полученной базой данных и состоит из следующих технологических операций:
- автоматический контроль с формированием протокола полученных автоматически и введенных интерактивно объектов на соответствие их классификатору и правилам ввода;
 - по протоколу, полученному на предыдущем этапе, интерактивно исправляются ошибочные объекты;
 - формируется топология объектов в базе данных путем установления пространственно-логических (ПЛС) и тополого-метрических (ТМС) связей;
 - осуществляется автоматизированный контроль всей введенной информации;
 - производится загрузка полученной информации в оригинальный обменный формат для долговременного хранения.
4. Если не все планы этажей данного здания введены, то переходим к вводу плана следующего этажа:
- изначально производится оценка отличий данного этажа от введенного первым. При отличии, не превышающем 10%, производится исключение (логическое вычитание) из файла штрихов, вводимого этажа файла штрихов этажа введенного первым; создается копия БД введенного первым этажа;
 - используя исходную растровую модель плана в штриховой форме, осуществляем ввод не введенных на предыдущем этаже объектов и добавление их в базу данных, полученную на предыдущем этапе. Затем из файла штрихов этажа, введенного первым, исключаются штрихи обрабатываемого этажа. По полученным штрихам из базы данных удаляются отсутствующие на данном этаже объекты;
 - для полученной базы данных выполняются все технологические операции, начиная с 3 технологического этапа;
 - если объективное отличие текущего этажа от этажа, введенного первым, превышает 10%, то его обработка проводится начиная со 2 технологического этапа.
5. После формирования баз данных всех этажей здания производится сопряжения (установление соответствия и запись ссылок на все полученные базы данных в поле мультимедиа соответствующего здания).
6. Базы данных поэтажных планов помещаются на хранение в архив. На этом обработка поэтажных планов здания заканчивается.
- Данная технология разработана и внедрена в рамках развития ГИС «Терра», созданной в НИИ ПМК (ныне Центр Информатики и Интеллектуальных Информационных Технологий ИТММ) ННГУ им. Н. И. Лобачевского.

Геометрическое конструирование на параметрической основе с элементами эвристического поиска*

М. В. Лагунова, Т. В. Мошкова, В. А. Тюрин
mvlnn@mail.ru|ng.forever.mtv@gmail.com|55555_73@mail.ru

Нижегородский государственный архитектурно-строительный университет

Рассмотрена проблема автоматизированного синтеза предварительной каркасной модели трёхмерного объекта по техническому чертежу, содержащему разрезы и сечения. Описанный процесс позволяет расширить возможности алгоритма синтеза электронной модели 3D объекта по данным технического чертежа, разрабатываемого нижегородской научной геометрической школой под руководством профессора Роткова С.И. Сложности формализации анализа изображений, составляющих содержание технического чертежа, накладывают ограничения на входную информацию алгоритмов, решающих задачу синтеза, поэтому работа по расширению классов входной информации является актуальной и востребованной.

Ключевые слова: электронная геометрическая модель изделия, каркасная модель, предварительная каркасная модель, технический чертеж, разрезы и сечения, автоматизированное проектирование.

Geometric construction on a parametric basis with elements of heuristic search*

M. V. Lagunova, T. V. Moshkova, V. A. Tyurina

Nizhny Novgorod State University of Architecture and Civil Engineering

The problem of automated synthesis of the preliminary wireframe model of a three-dimensional object is considered according to a technical drawing containing sections and sections. The described process makes it possible to expand the possibilities of the algorithm for synthesizing the electronic model of a 3D object according to the technical drawing developed by the Nizhny Novgorod scientific geometric school under the direction of Professor Rotkov S.I. The difficulties in formalizing the analysis of images that make up the contents of a technical drawing impose serious limitations on the input information of algorithms that solve the synthesis problem, therefore the work on expanding the classes of input information is very relevant.

Keywords: electronic geometric model of the product, wireframe model, preliminary wireframe model, technical drawing, sections and sections, computer-aided design.

1. Введение

Развитие компьютерных технологий проектирования, расширение возможностей САД-систем оказывает существенное влияние на компоненты профессиональной деятельности в строительстве, архитектуре, дизайне и многих других областях. Развитое конструктивно-геометрическое мышление становится более востребованным и меняет требования к геометро-графической подготовке специалистов инженерно-строительного профиля.

В процессе геометро-графической подготовки традиционным является выполнение графических работ, призванных в соответствии с целями обучения развить конструктивно-геометрическое мышление, способности к анализу и синтезу простран-

ственных форм и отношений, изучить способы конструирования различных геометрических объектов, способы получения их чертежей. В традиционном подходе к комплектованию графических работ на разных этапах геометро-графической подготовки студентов строительных направлений доминирует автономность без достаточных логико-содержательных связей не только между дисциплинами одного цикла, но и между работами внутри одной дисциплины, что связано в большей степени уменьшением академического объема дисциплин при сохранившейся дидактической насыщенности [1, 2, 3, 4]. Графические работы ориентированы преимущественно на нормативную схему процесса графической деятельности, определяющую последовательность разработки графического объекта на основе типологии конструкторских задач и соблюдении стандартных приемов и опера-

Работа выполнена и опубликована при финансовой поддержке РФФИ, гранты 17-07-00543, 18-07-20030.

ций, регламентируемых системой ЕСКД и особенностями функционала САД-систем.

Параметрический подход [5] к конструированию не находит еще широкой реализации в содержательном наполнении практических графических работ, в то время как одной из важнейших задач, связанных с повышением качества подготовки специалистов и соответственно проектной деятельности в целом, является переход на параметрические цифровые модели. Представляется актуальным исследование возможностей организации графической деятельности студентов, обеспечивающей развитие таких параметров конструктивно-геометрического мышления как целостность восприятия, субъективность, многозначность, динамичность сформированного геометрического образа в ходе решения конструктивных задач на параметрической основе на разном уровне в различных технологиях исполнения.

2. Постановка задачи

Преодоление противоречия между алгоритмическими способами переработки данных в информационных системах компьютерной графики и неалгоритмическими способами конструктивно-геометрического мышления обучаемых предполагает определенную перестройку мышления обучаемых при рассмотрении процессов параметрического конструирования на основе усиления теоретической компоненты содержания и поэтапного обучения с постепенным усложнением рассматриваемых процессов и структур.

Предлагаемая на кафедре Инженерной геометрии, компьютерной графики и автоматизированного проектирования ФГБОУ ВО «Нижегородский государственный архитектурно-строительный университет» методика геометро-графической подготовки включает конструктивно-геометрические задачи на основе параметризации в рамках изучения дисциплин «Начертательная геометрия», «Компьютерная графика» при обучении бакалавров направления

08.03.01 Строительство, 07.03.01 Архитектура, 07.03.09 Дизайн архитектурной среды. Это обусловлено тем, что оптимизация конструктивно-геометрических задач и временных затрат проектной деятельности лежит в фазах конструирования, в процессе которых идет поиск принципов конструирования (определение геометрии элементов), определение расположения элементов (определение вариантов геометрии) и детализовки, в процессе которой идет определение окончательного решения, оптимизация зон конструирования, собственно само выполнение чертежей.

Предлагаемая методика на разных этапах обучения включает необходимость разработки серии плоских контуров, представляющих несколько вариантов решения конструкторской задачи по заданным критериям. Данная задача относится, по классификации Полозова В.С., к неформальным задачам [5]. Однако решение таких задач возможно с применением эвристических приемов. Это позволяет в качестве методологического подхода использовать обобщенный эвристический алгоритм [6].

Логически обоснованный процесс решения предлагаемой конструкторской задачи многофункционален:

- усвоение общих понятий геометрического описания предметного мира, развитие навыков анализа и синтеза плоских форм и их отношений;
- осмысление сущности, цели, задач, функций геометрического конструирования;
- осмысление теории параметризации, усвоение основных понятий, принципов;
- обоснование, выявление этапов решения задач геометрического конструирования на параметрической основе как проектно-творческой деятельности;
- освоение средств и практических графических приемов решения задачи конструирования в ручной технике;
- уяснение требований и развитие навыков оформления чертежей в соответствии с нормативами оформления графических работ;

- оценку различных аспектов деятельности геометрического конструирования, включая оценочный, презентационный и другие аспекты.

3. Реализация этапов геометрического конструирования плоского контура на параметрической основе

В соответствии с общим подходом нами выделены следующие этапы учебного геометрического конструирования плоского контура на параметрической основе:

1. Подготовительный этап.
2. Этап творческого поиска.
3. Этап вариантного моделирования.
4. Этап творческой разработки.
5. Этап параметризации.
6. Исполнительный этап выбранного технического решения.
7. Этап ретроспективной оценки и методический разбор.

На подготовительном этапе происходит теоретическая подготовка и сбор необходимой информации, необходимой для конструирования. Элементы эвристического поиска должны опираться на упорядоченный запас знаний, в том числе усвоение общей теории параметризации, усвоение основных понятий, условий и ограничений, усвоение знаний о приемах построения элементарных геометрических фигур и их параметрах, алгоритмах комплексирования элементарных фигур (алгоритмы пересечения, сопряжения, касания и др.). На этом этапе студентам предлагается знакомство с стандартными решениями — эталонами графической модели. Демонстрируются варианты решения задач студентами-предшественниками в различных технологиях исполнения: ручной и автоматизированной. Знакомство с эталонами обеспечивает возможность критической оценки, обеспечивает создание ориентировочной основы деятельности и стимулирует к творческому поиску. Предметом конструирования являются контуры различного назначения, обладающие гармоничной (содержательной, технологической, эстетической)

формой, отражающей в визуально воспринимаемых признаках ценность объекта проектирования по заданным критериям. Параметрическая конструкция позволяет решать задачи исследования процесса формообразования и локального изменения формы. Учет, наложение или снятие геометрических условий позволяет исследовать параметрическую модель. На рисунке 1 в качестве ориентировочной основы действий показан процесс конструирования плоского контура с использованием параметрических возможностей системы геометрического моделирования КОМПАС-3D. На простом примере показан подсчет минимального (но достаточного для однозначного воспроизведения фигуры) числа параметров контура, проведенный с одновременным построением изображения и наложением связей на геометрические элементы, составляющие контур. После учета геометрических условий выявлены семь независимых формальных параметров, определяющих форму и положение данного плоского контура в выбранной системе координат. Им поставлены в соответствие семь переменных, которым присвоены фактические значения, позволившие получить приведенное на рис. 1 изображение.

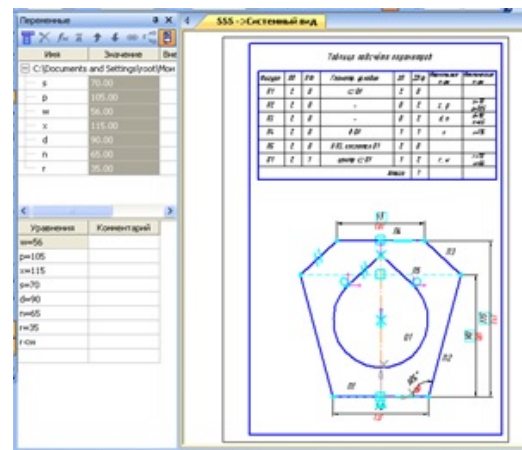


Рис. 1: Пример изображения, полученного в результате реализации этапов геометрического конструирования плоского контура

Наложённые геометрические связи вызывают изменения в количестве степеней свободы движения характерных точек контура

ра. Например, после наложения на центр окружности $O1$ требования принадлежности оси OY , ее центр имеет возможность двигаться только вдоль этой оси (т.е. имеет одну степень свободы). Связи и степени свободы показываются на чертеже в системе КОМПАС-3D специальными обозначениями. Если независимые переменные не фиксируются, а используются как «информационные», то, меняя положение подвижных точек контура (или вводя фактические значения переменных), можно изменять его форму, не нарушая наложенных связей. Изменение фактических значений переменных позволяет выбрать наиболее оптимальную конфигурацию контура, в зависимости от требований к нему в каждом конкретном случае. На рис. 2 показаны четыре варианта описанного выше параметрического контура.

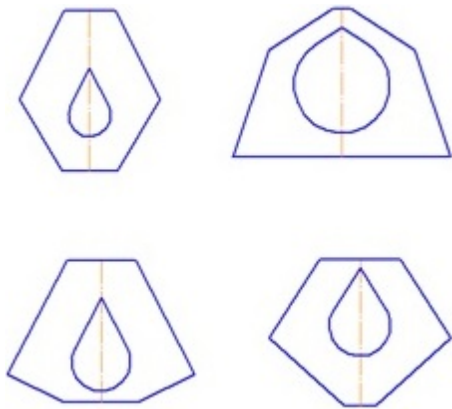


Рис. 2: Четыре варианта параметрического контура

На этапе творческого поиска происходит выдвижение исходной авторской идеи конструируемого объекта с последующей конкретизацией. Эскизная реализация, иллюстрирующая поиск идеи с дальнейшей деконпозицией и аппроксимацией контура в соответствии с заданными ограничениями проходит коллективное обсуждение с преподавателем и другими студентами. Это позволяет развить наиболее содержательные творческие потенции разработчика, выявить перспективные направления развития принципиального замысла.

На этапе вариантного моделирования дальнейшее конструирование подчинено обеспечению взаимоположения и организации функциональных зон и узлов контура с точки зрения наложенных геометрических условий и структурных и алгоритмических ограничений. Особое значение при этом приобретают операции преобразований, способствующие изменению формы (масштабирование, подобие) структуры либо положения объекта (повтор, перенос, перемещение, сдвиг), копирование с перемещением по заданному закону, зеркальное отображение, поворот, определение геометрических областей пересечения, вычитание, объединение, отсоединение, присоединение, удаление).

На этапе творческой разработки контура происходит согласование составляющих геометрической модели, определяющих форму и положение элементов контура с учетом выбранных геометрических условий, выбор системы параметризации, введение параметров (независимых переменных) и составление параметрического расчета, представленного в табличной форме.

Большое обучающее значение при этом имеет необходимость перекодирования графической информации в символьную и обратно. Активизация понятийного мышления, укрепляет сущностные связи между геометрическими примитивами и позволяет выбрать рациональный порядок выполнения контура с целью минимизации параметрического числа и реализации заложенных графических алгоритмов. Далее происходит реализация параметрического числа размерной сетью в соответствии с ГОСТ 2.307-2011 с назначением фактических значений переменных.

На этапе исполнительного этапа конструирования происходит графическое исполнение контура в соответствии с требованиями, регламентированными конструкторскими правилами и нормами (ЕСКД, алгоритмы геометрических построений). Этот процесс сопровождается развитием навыков работы с информационными ис-

точниками благодаря доступу к единой информационно-образовательной среде вуза. На этапе ретроспективной оценки и методического разбора происходит прием, оценка и защита конструкторского решения. Цель методического разбора — оценка оптимальности предложенного решения, выявление возможностей уменьшения параметрического числа контура при сохранении его геометрической концепции.

На следующем уровне обучения в рамках дисциплины «Компьютерная графика» геометрическое конструирование включает следующие этапы с новым содержательным и технологическим наполнением:

1. Подготовительный этап.
2. Исполнительный этап выбранного технического решения.
3. Этап параметризации.
4. Этап вариантного моделирования.
5. Этап ретроспективной оценки и методический разбор.

Подготовительный этап включает ознакомление с инструментарием САД-системы. Исполнительный этап предполагает практическое освоение инструментария при исполнении контура в САД-системе. Этап параметризации наполняется новыми технологическими приемами, обусловленными возможностями применяемой системы. Этап вариантного моделирования параметризованного контура позволяет решать задачи исследования процесса формообразования и локального изменения формы. Творческие практические задания по конструированию авторских плоских контуров в рамках изучения компьютерного практикума на этапе изучения дисциплины «Компьютерная графика» способствуют более осмысленному освоению автоматизированного параметрического конструирования, поскольку используются авторские контуры, разработанные на первом этапе обучения (рис. 3,4).

Возможности компьютерной технологии способствуют исследованию геометрических свойств модели, при этом наглядность

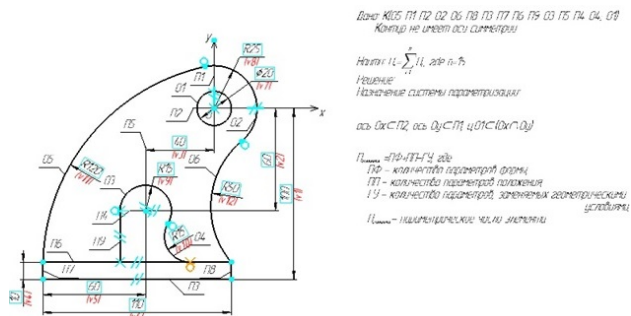


Рис. 3: Авторские контуры, разработанные на первом этапе обучения

Таблица подсчета параметров

№	Условие	П1	П2	Геометрические условия	ΣГУ	ΣП	Размер, мм
1	P11	2	0	P11 = Oy	2	0	-
2	P12	2	0	P12 = Ox	2	0	-
3	P13	2	0	P13 = Ox	1	1	v1=100
4	P14	2	0	P14 = P2	1	1	v2=60
5	P15	2	0	P15 = Oy	1	1	v3=40
6	P16	2	0	P16 = P3	1	1	v4=10
7	P17	2	0	P17 = P5	1	1	v5=60
8	P18	2	0	P18 = P7	1	1	v6=110
9	01	2	1	u.01 = P11 - P12	1	1	v7=±20
10	02	2	1	u.02 = u.01	1	1	v8=R25
11	03	2	1	u.03 = P16 - P15	1	1	v9=R15
12	P19	2	0	P19 = O. 03, P19 = P5	2	0	-
13	04	2	1	04 = O. P16, 04 = O. 03	2	1	v10=R15
14	05	2	1	05 = O. 02, 05 = P16 - P17	2	1	v11=R120
15	06	2	1	06 = O. 02, 06 = P18 - P16	2	1	v12=R50
Параметрическое число контура					12	12	

Рис. 4: Таблица расчета параметров авторских контуров

позволяет быстрее и качественнее принять оптимальное решение при моделировании сложных форм. Учет, наложение или снятие геометрических условий позволяет исследовать параметрическую модель. Обучаемый переживает эмоциональный подъем в те моменты, когда приходит к успешному освоению очередного набора навыков формообразования, демонстрирующего на экране компьютера наглядное динамичное образование смоделированной в сознании студента композиции.

4. Заключение

Процесс обучения с использованием эвристических элементов позволяет перейти от неформальных задач к построению формальных моделей геометрического моделирования с помощью САД-систем. Предложенная методика обучения, реализованная

на практике, позволяет выделить ее особенности:

- приоритет самостоятельности обучения предполагает самостоятельное определение исходной идеи контура, вариантное моделирование контура, выбор геометрических условий и отношений входящих примитивов, обеспечивающий реализацию ценностного подхода в обучении, повышающий уровень личностной мотивации обучаемых;
- совместной деятельности включение в процесс планирования, оценивания результатов обучения обучаемых и обучающихся, перенос акцента в работе преподавателя в сторону консультативно-координирующих функций управления познавательной деятельности обучаемых;
- индивидуализации обучаемый сам выбирает уровень сложности конструирования объекта в соответствии со своими потребностями и способностями;
- актуализации результатов применение полученных знаний и умений, востребованных на новом этапе обучения внутри дисциплины и дисциплинарного цикла, формирующую у обучаемых основу действий высокого уровня обобщения.
- устранения параллелизма, дублирования и информационных перегрузок обучающихся.

В ходе педагогического наблюдения за деятельностью студентов в процессе лабораторного практикума выделены группы личностных факторов, которые наряду с необходимой базой знаний и умений традиционного проектирования, влияют на успешность и интерес к решению задач формирования с помощью информационных технологий:

- стремление к получению новой информации;
- организованность и настойчивость в поиске новой информации;
- стремление к принятию независимых авторских решений;

- ведущий способ взаимодействия с объектами – структурирование материала.

Исследование возможностей организации графической деятельности студентов на параметрической основе позволило обеспечить развитие основных параметров конструктивно-геометрического мышления как целостность восприятия, субъективность, многозначность, динамичность сформированного геометрического образа в ходе решения конструктивных задач на параметрической основе на разном уровне в различных технологиях исполнения.

Об авторах

Лагунова Марина Викторовна, д.п.н., профессор, профессор кафедры Инженерной геометрии, компьютерной графики и автоматизированного проектирования Нижегородского государственного архитектурно-строительного университета
mvlmn@mail.ru

Мошкова Татьяна Владимировна, к.т.н., профессор кафедры Инженерной геометрии, компьютерной графики и автоматизированного проектирования Нижегородского государственного архитектурно-строительного университета
ng.forever.mtv@gmail.com

Тюрина Валерия Александровна, к.т.н., профессор кафедры Инженерной геометрии, компьютерной графики и автоматизированного проектирования Нижегородского государственного архитектурно-строительного университета
55555_73@mail.ru

Литература

- [1] Рукавишников В.А. Геометро-графическая подготовка инженера: время реформ // Высшее образование в России, №5, 2008. – С. 132-136.
- [2] Амирджанова И.Ю., Виткалов В.Г. Современное состояние развития геометро-графической культуры и компетентности будущих специалистов // Вектор науки ТГУ. 2015. № 2 (32-2). С. 26-31.
- [3] Юматова Э.Г. Системно-функциональный анализ проблем, противоречий и способов их разрешения в геометро-графической подготовке студентов технических вузов / Современные наукоемкие технологии. 2016. № 8-1. С. 174-178.

- [4] Лагунова М.В., Елисеева М.Ю. Формирование готовности студентов технических вузов к конструкторско-графической деятельности: Монография. – Нижний Новгород: ВГИПУ, 2008. – 115 с.
- [5] Полозов В.С. Начертательная геометрия (информационно-параметрический подход): Уч. пособ. – Нижний Новгород, ННГАСУ, 2000 – 61 с.
- [6] Автоматизация поискового конструирования. Под ред. А. И. Половинкина. – М.: Радио и связь, 1981. – 344 с.