

# Оглавление

<b>Аппаратная база и общая теория программирования</b>	<b>11</b>
Экспериментальное сравнение средств обнаружения клонов <i>Бульчѳв П.Е., Шалимов А.В., Коваленко Д.С.</i> . . . . .	11
Апробация технологии тестирования UniTESK на реализациях паттернов управления взаимодействием в AJAX приложениях <i>Герлиц Евгений Анатольевич</i> . . . . .	12
О методах анализа полноты требований <i>Гингина Виктория Владимировна</i> . . . . .	13
Исследование и разработка методов повышения эффективности гибридных параллельных программ для высокопроизводительных систем <i>Глазкова Екатерина Алексеевна</i> . . . . .	15
О разметке области стека в задаче обратной инженерии <i>Деревенец Егор Олегович</i> . . . . .	16
Система справедливого планирования задач пользователя, на основе генетического алгоритма и унифицированного запуска задач через веб-интерфейс на разных вычислительных комплексах. <i>Князев Николай Александрович</i> . . . . .	17
Библиотека Libjit для динамической компиляции <i>Кононенко Кирилл Владимирович</i> . . . . .	19
Определение стратегии вытеснения Pseudo-LRU на ветвях бинарного дерева <i>Корнышин Евгений Валерьевич</i> . . . . .	20
Решение задачи N тел на гибридных архитектурах с использованием библиотеки ttgLib <i>Кривов М.А., Гризан С.А.</i> . . . . .	22
Автоматическое динамическое распараллеливание вычислений для гетерогенных многопроцессорных компьютерных систем <i>Левченко Роман Игоревич</i> . . . . .	23
Применение генетических алгоритмов в задаче проектирования отказоустойчивых вычислительных систем <i>Пашков В.Н., Волжанов Д.Ю.</i> . . . . .	24

Тестирование соответствия почтовых протоколов сети Интернет <i>Тугаенко Анастасия Николаевна</i> . . . . .	25
<b>Дискретная математика и математическая кибернетика</b>	<b>28</b>
Легкотестируемые схемы для дизъюнкции <i>Беджанова Светлана Руслановна</i> . . . . .	28
Свойства равновесий в коалиционных играх с неравномерным распределением агентов <i>Вартанов Сергей Александрович</i> . . . . .	29
О семействах точных вершинных $k$ -расширений графов при $k > 1$ <i>Долгов Александр Алексеевич</i> . . . . .	30
Репрезентативность исходов аукциона функций предложений с учетом неопределенности спроса <i>Долматова Марина Станиславовна</i> . . . . .	32
Получение «справедливых» устойчивых дележей в задачах кооперативных игр <i>Клименко Оксана Александровна</i> . . . . .	33
Инъективная $L(2, 1)$ -раскраска расщепляемых неразложимых униграфов <i>Максимович Олег Валерьевич</i> . . . . .	34
Алгоритмы решения систем булевых уравнений с использованием ассоциативных принципов обработки информации. <i>Мелузов Антон Сергеевич</i> . . . . .	35
Расшифровка обобщенных псевдо-булевских монотонных функций <i>Осокин Виктор Владимирович</i> . . . . .	36
Равновесия двухэтапного рынка со случайным исходом на спотовом рынке. <i>Шарикова Агата Андреевна</i> . . . . .	38
<b>Дифференциальные уравнения, динамические системы и оптимальное управление</b>	<b>40</b>
О задаче Коши для нелинейного уравнения соболевского типа с переменным коэффициентом <i>Аристов Анатолий Игоревич</i> . . . . .	40
Вероятностная модель межрегиональных отношений в контексте поставок природного газа <i>Борисова Анастасия Викторовна</i> . . . . .	41
Оценивание рисков в экологических сетях <i>Вещинская Виктория Валерьевна</i> . . . . .	42
Аппроксимация множеств достижимости управляемых систем в среде Matlab <i>Винников Евгений Владимирович</i> . . . . .	44

Применение операторной обратной связи для экспоненциальной стабилизации билинейных систем <i>Гончаров Олег Игоревич</i> . . . . .	46
Метод имитационного моделирования групповой динамики толпы <i>Гребенников Роман Владимирович</i> . . . . .	47
Синтез оптимального управления в математической модели терапии вирусных инфекций <i>Зайчик Светлана Юрьевна</i> . . . . .	48
Верхние оценки порядка стабилизатора для ММО систем <i>Капалин Иван Владимирович</i> . . . . .	49
Динамическое программирование в линейных системах с состояниями в виде распределений <i>Мазуренко Станислав Сергеевич</i> . . . . .	51
Оптимизация биотехнологического процесса с использованием методов математического моделирования <i>Плотникова Юлия Валентиновна</i> . . . . .	52
О внутренних и внешних оценках информационных множеств задачи гарантированного оценивания <i>Прозоров Илья Сергеевич</i> . . . . .	53
Программное приложение для параметрической идентификации динамических моделей <i>Семёнов Михаил Игоревич</i> . . . . .	54
<b>Компьютерная графика</b>	<b>56</b>
Анализ изображений на основе характеристик краев и формы объектов <i>Жариков Александр Леонидович</i> . . . . .	56
Система трёхмерной реконструкции объектов и калибровки проекторов с использованием структурированного света <i>Зейналов Рамиз Шакирович</i> . . . . .	57
Реконструкция текстур сильно загороженных фасадов зданий с использованием нескольких изображений. <i>Косонов В.А., Копушин В.С.</i> . . . . .	59
Алгоритм автоматического совмещения аэрофотоснимков с малым перекрытием <i>Кривовязь Г.Р., Мизотин М.М.</i> . . . . .	60
Система стеганографии для видео последовательностей <i>Куликов Дмитрий Леонидович</i> . . . . .	61
Физически аккуратное моделирование отражательной способности материалов поверхностей плоских объектов по фотоизображениям <i>Лебедев А.С., Ильин А.А.</i> . . . . .	63

Алгоритм устранения эффекта Гиббса при восстановлении сжатого мультипликационного видео <i>Моисейцев Алексей Борисович</i> . . . . .	64
Система построения трехмерных моделей зданий по поэтажным планам <i>Петрушкина Анастасия Александровна</i> . . . . .	65
Восстановление параметрических моделей отражательных свойств материалов по двум фотоизображениям <i>Синяевский В.А., Ильин А.А.</i> . . . . .	67
Синтез HDR изображений при помощи технологии CUDA <i>Скрябин Виктор Сергеевич</i> . . . . .	68
Графическая модель для оценки геометрии городских сцен по одному изображению <i>Третьяк Елена Викторовна</i> . . . . .	69
Конвейер калибровки и текстурирования модели головы по одной фотографии <i>Федюков Максим Александрович</i> . . . . .	71
Адаптивная трассировка путей на графических процессорах <i>Фролов Владимир Александрович</i> . . . . .	72
Автоматический анализ и обработка снимков пищеварительного тракта, сделанных в процессе капсульного эндоскопического исследования <i>Хабидуллин Т.Р., Михайлов Д.М.</i> . . . . .	72
О подходе к объективной оценке качества цветных цифровых изображений <i>Хорунжий Михаил Дмитриевич</i> . . . . .	74
Ускорение обработки видеоданных с применением современных графических процессоров <i>Цветков Илья Игоревич</i> . . . . .	76
Калибровка камеры по нескольким шаблонам <i>Четвериков Никита Сергеевич</i> . . . . .	77
Классификация трёхмерных облаков точек с помощью неассоциативных Марковских сетей <i>Шаповалов Роман Викторович</i> . . . . .	78
<b>Машинное обучение и анализ данных</b>	<b>80</b>
Автоматическое выявление терминологических вариантов в русскоязычных текстах <i>Антонов В.Ю., Ефремова Н.Э.</i> . . . . .	80
Применение нейронных сетей для построения механических торговых систем. <i>Глекин Александр Олегович</i> . . . . .	81

Классификация прецедентов, образующих связные множества, на примере данных лазерной локации <i>Гусейнов Алексей Тимурович</i> . . . . .	82
Разрешимость задач нечёткой классификации элементов точечных конфигураций <i>Дорофеев Николай Юрьевич</i> . . . . .	83
Условия корректности линейного замыкания модели АВО <i>Иофина Галина Владимировна</i> . . . . .	85
Применение генетических алгоритмов для выделения событий в информационном потоке. <i>Карпов Илья Андреевич</i> . . . . .	86
Разделение системы точек на подмножества с невырожденными матрицами попарных расстояний <i>Карпович Павел Алексеевич</i> . . . . .	87
Порождение и отбор признаков при восстановлении линейной регрессии <i>Крымова Екатерина Александровна</i> . . . . .	88
Поиск взаимозаменяемых товаров <i>Ломакин Василий Дмитриевич</i> . . . . .	89
Экспериментальное исследование алгоритма отбора подмножеств взаимосвязанных признаков в задаче обучения распознаванию образов <i>Маслакова Светлана Николаевна</i> . . . . .	91
О метрическом аналоге метода главных компонент <i>Найденев Никита Анатольевич</i> . . . . .	92
Инкрементные методы в коллаборативной фильтрации <i>Полежаева Елена Андреевна</i> . . . . .	93
Поиск похожих пользователей в социальных сетях методами коллаборативной фильтрации <i>Пустовойтов Никита Юрьевич</i> . . . . .	95
Оценки вероятности переобучения, зависящие от графа расслоения и связности семейства алгоритмов <i>Решетняк Илья Михайлович</i> . . . . .	96
Построение регрессионных моделей прогноза временных рядов <i>Сологуб Роман Аркадьевич</i> . . . . .	97
Точная оценка вероятности переобучения для одного специального семейства алгоритмов <i>Толстихин Илья Олегович</i> . . . . .	98
Метод извлечения оценочных слов из текстов <i>Четвёркин Илья Игоревич</i> . . . . .	99
Самообучаемая система коррекции L <sup>A</sup> T <sub>E</sub> X-документов <i>Чувилин Кирилл Владимирович</i> . . . . .	101

Моделирование поведения группы агентов с помощью взаимосвязанных скрытых марковских моделей <i>Юданов Анатолий Александрович</i> . . . . .	102
Применение генетического программирования для обучения решающих деревьев в контексте процедуры бустинга. <i>Янгель Борис Константинович</i> . . . . .	103
<b>Парадигмы программирования и интеллектуальная обработка данных</b>	<b>105</b>
Онтологический подход для репрезентации интенциональной парадигмы языка MSC <i>Вдовенко Татьяна Владимировна</i> . . . . .	105
Разработка системы интеллектуального анализа данных для оптимизации процесса электронного документооборота высшего учебного заведения <i>Гладких Николай Александрович</i> . . . . .	106
Поиск фразеологизмов в тексте <i>Грибанов Д.В., Смычѣк Е.М.</i> . . . . .	108
Оценка времени анализа событий системой обнаружения и предотвращения сетевых атак <i>Казачкин Дмитрий Сергеевич</i> . . . . .	109
Автоматическая генерация тестов по комментариям к программному коду <i>Носков Алексей Анатольевич</i> . . . . .	111
Интеллектуальная информационная система атрибуции художественных текстов <i>Попов Никита Евгеньевич</i> . . . . .	112
Итерационная модель нейросетевой ассоциативной памяти <i>Прокопович Григорий Александрович</i> . . . . .	113
Способ моделирования XML-данных выражениями языка C++ <i>Стружков Илья Викторович</i> . . . . .	115
Система создания и управления мультимедийными документами Fuzzle CMS <i>Январев Владислав Игоревич</i> . . . . .	116
<b>Теория вероятностей и математическая статистика</b>	<b>118</b>
Асимптотические оценки абсолютной постоянной в неравенстве Берри-Эссеена для распределений, не имеющих третьего момента <i>Гапонова Маргарита Олеговна</i> . . . . .	118
Уточнение неравенства Каца-Берри-Эссеена <i>Григорьева Мария Евгеньевна</i> . . . . .	120

Исследование применимости моделей оценки надёжности для разработки программного обеспечения с открытым исходным кодом. <i>Зорин Даниил Александрович</i> . . . . .	121
Классификация многоканальных дистанционных изображений с использованием марковских случайных полей и копул <i>Крылов Владимир Андреевич</i> . . . . .	122
Асимптотическая нормальность оценок риска пороговой обработки вейвлет-коэффициентов <i>Маркин Артём Васильевич</i> . . . . .	123
Численный поиск нижней оценки для абсолютной константы в неравенстве Берри-Эссеена для пуассоновских случайных сумм. <i>Нефедова Юлия Сергеевна</i> . . . . .	124
Оптимальный выбор числа допустимых ошибок первого рода при множественной проверке гипотез <i>Рябенко Евгений Алексеевич</i> . . . . .	126
Вейвлет-обработка цветных теневых картин аэродинамического эксперимента <i>Хазиахметов Максим Шамилович</i> . . . . .	127
Метод разумных целей для многокритериальных задач выбора с вероятностными данными <i>Холмов Алексей Владимирович</i> . . . . .	130
Распределение вероятностей локального наклона фазового рельефа для задачи развертки фазы в космической радиолокационной топографической интерферометрии <i>Шувалов Роман Игоревич</i> . . . . .	131
<b>Численные методы и математическое моделирование</b>	<b>133</b>
Аппроксимация решения гиперсингулярного интегрального уравнения Прандтля тензорным поездом <i>Андрюченко Назар Юрьевич</i> . . . . .	133
Моделирование двумерных нестационарных газодинамических процессов методом ленточно-адаптивных сеток <i>Белевский Владислав Викторович</i> . . . . .	135
Численное решение трехмерной задачи двухфазного течения жидкостей в пористой среде <i>Борисов Виктор Светославович</i> . . . . .	136
Сплайн-интегрирование систем обыкновенных дифференциальных уравнений с запаздыванием <i>Власов Сергей Александрович</i> . . . . .	139
Монотонный метод конечных объемов для задачи конвекции-диффузии на сетках с многогранными ячейками <i>Данилов Александр Анатольевич</i> . . . . .	140

Построение стохастических моделей в задачах среднесрочного прогнозирования солнечной активности <i>Кудрявцев Ю.В., Шмалько Р.А.</i> . . . . .	140
Блочный тензорный метод типа сопряжённых градиентов для минимизации отношения Рэлея в QTT-формате <i>Лебедева Ольга Сергеевна</i> . . . . .	142
Математическое моделирование процесса заполнения лазерных мишеней ядерным топливом <i>Малинина Елена Александровна</i> . . . . .	143
Анализ применимости конечно-элементных моделей для решения нестационарных задач сейсродинамики и исследования эффективности применения горизонтальных сейсмических барьеров. <i>Нафасов А.Э., Кузнецова М.Н.</i> . . . . .	144
Технология моделирования течений со свободной поверхностью в реалистичных сценах. <i>Никитин Кирилл Дмитриевич</i> . . . . .	146
Сравнительный анализ эффективности некоторых преобуславлявателей. <i>Никольский Илья Михайлович</i> . . . . .	147
Ньютоновские методы для задач оптимизации с комплементарными ограничениями <i>Погосян Артур Левоневич</i> . . . . .	148
Алгоритм построения сильно различающихся решений некоторого класса обратных задач <i>Сучков Егор Петрович</i> . . . . .	149
Моделирование развития инфекции с использованием агентного подхода <i>Улыбин Андрей Владимирович</i> . . . . .	150
Вычисления из первых принципов пути реакции изомеризации в молекулярной системе методом метадинамики. <i>Шумкин Георгий Николаевич</i> . . . . .	151



# Аппаратная база и общая теория программирования

## Экспериментальное сравнение средств обнаружения клонов

*Булычёв П.Е., Шалимов А.В., Коваленко Д.С.*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: peter.bulychev@gmail.com*

Клоном называют два схожих фрагмента кода. Часто фрагменты, составляющие клон, реализуют одну и ту же функциональность, поэтому при изменении одного фрагмента необходимо изменить и другой. Поэтому присутствие клонов в исходном коде увеличивает стоимость его поддержки, и средства обнаружения клонов могут использоваться в процессе улучшения качества кода путём рефакторинга.

Целью данной работы является экспериментальное сравнение средства обнаружения клонов Clone Digger [1], разработанное одним из авторов, с существующими средствами CloneDR [2] и Dup [3]. Методика данного эксперимента взята из работы Бэллона [4]. Мы выбрали CloneDR для сравнения, поскольку в эксперименте Бэллона оно имело высокий показатель точности поиска, и алгоритм работы CloneDR схож с алгоритмом работы Clone Digger. Средство Dup было выбрано потому, что в эксперименте Беллона оно имело наилучший показатель полноты поиска.

Сравниваемые средства были запущены на проекте Eclipse JDT Core [5], написанном на языке Java и состоящем из 146 тысяч строк кода. Затем 2% от 31500 найденных клонов были проанализированы вручную и было сформировано эталонное множество, в которое попали только те клоны, от которых, по мнению экспертов (авторов данной заметки), можно избавиться при помощи рефакторинга. При этом эксперты не знали, каким именно средством был обнаружен тот или иной клон. Для каждого из трёх средств была вычислена полнота поиска (процент от числа клонов из эталонного множе-

ства, обнаруженных данным средством) и точность поиска (процент от числа клонов, обнаруженных выбранным средством, попавших в эталонное множество). Оказалось, что средство Clone Digger превосходит два других средства как по полноте, так и по точности поиска; средство Dup превосходит средство CloneDR по полноте поиска, но имеет меньшую точность.

В докладе будут представлены детали и результаты эксперимента.

### Литература

1. Bulychev P., Minea M. Duplicate code detection using anti-unification // SYRCoSE, 2008.
2. Baxter I. D., Yahin A., Moura L., Anna M. S., Bier L. Clone Detection Using Abstract Syntax Trees // ICSM, 1998.
3. Baker B. S. A program for identifying duplicated code // CSS, 1992.
4. Bellon S., Koschke R., Antoniol G., Krinke J., Merlo E.. Comparison and Evaluation of Clone Detection Tools // IEEE Transactions on Software Engineering, 2007.
5. Eclipse JDT Core: <http://eclipse.org/jdt/core>

## Апробация технологии тестирования UniTESK на реализациях паттернов управления взаимодействием в AJAX приложениях

*Герлиц Евгений Анатольевич*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,*

*ВМК, Москва, Россия*

*email: [gerlits@ispras.ru](mailto:gerlits@ispras.ru)*

Классические веб-приложения, основная функциональность которых реализована на стороне сервера, значительно уступают в интерактивности веб-приложениям с асинхронным интерфейсом, построенным с применением технологии AJAX (Asynchronous JavaScript and XML). Между тем интерактивность интерфейсов AJAX приложений достигается за счет значительного усложнения их реализации. А именно, между браузером и веб-сервером появляется промежуточный уровень, в задачи которого входит управление событийно ориентированным взаимодействием на стороне клиента, управление коммуникацией между клиентом и сервером, а также модификация объектной модели документа.

Исследование существующих методик тестирования AJAX приложений показало, что они не нацелены на выявление ошибок в управлении взаимодействием на стороне клиента и между клиентом и сервером. Для тестирования

управления взаимодействием в AJAX приложениях мы предлагаем использовать технологию UniTESK и реализованный в ней подход к тестированию систем с асинхронным интерфейсом.

Ранее в работе [1] удалось продемонстрировать возможность применения технологии UniTESK к тестированию параллельных асинхронных взаимодействий в AJAX приложениях, однако апробация тогда была проведена лишь на двух AJAX приложениях. В данной работе для обоснования применимости технологии UniTESK к тестированию управления взаимодействием в AJAX приложениях прежде всего выделены существующие паттерны управления взаимодействием в AJAX приложениях, которые по сути своей описывают применяемые разработчиками схемы организации взаимодействия в AJAX приложениях. Для каждого паттерна разработано AJAX приложение, реализующее схему управления взаимодействием, определяемую паттерном, а для каждого AJAX приложения в свою очередь разработан тестовый набор, нацеленный на проверку схемы управления взаимодействием, реализованной в AJAX приложении.

В процессе тестирования в реализацию схемы управления взаимодействием AJAX приложений намеренно вносились ошибки, которые были успешно выявлены, что зафиксировано в тестовых отчетах. Испытания в целом подтвердили применимость технологии UniTESK к тестированию управления взаимодействием в AJAX приложениях. Между тем при разработке тестовых наборов для AJAX приложений приходится преодолевать ограничения в применении механизма сериализации отложенных реакций, а создание тестов представляет собой довольно трудоемкий процесс.

## Литература

1. Герлиц Е.А. Функциональное тестирование AJAX приложений // Сборник научных трудов Научно-практической конференции «Актуальные проблемы программной инженерии». М.: Издательство МЭСИ, 2009. С. 217-221

## О методах анализа полноты требований

*Гингина Виктория Владимировна*

*Аспирант*

*ИСП РАН, ИСП РАН, Москва, Россия*

*email: vgingina@ispras.ru*

Важную роль в разработке информационной системы играют требования. От качества описания требований зависит и этап разработки системы, и этап ее верификации. Поэтому существует необходимость описывать требования как можно более точно и корректно. Одним из свойств, определяющих качество составления требований, является полнота.

Достижение полноты требований обеспечит уменьшение количества доработок и изменений при разработке, снизит уровень проектных рисков, повы-

сит интероперабельность систем, сократит число ошибок, обнаруживаемых в процессе тестирования.

Критерий полноты требований можно разложить на две составляющие [1]. Полные требования должны описывать:

- во-первых, все, что заказчик хочет получить от системы;
- во-вторых, поведение системы во всех логически возможных ситуациях.

И если полнота в плане первого аспекта остается на откуп стороне-заказчику, то во втором аспекте анализ полноты может осуществляться специалистом по полноте требований.

С целью выявления источников неполноты требований и методов нахождения неполных требований ведется изучение проектной документации системы, содержащих требования стандартов, изучение уже выделенных требований, тестов, отчетов о тестировании и документов по управлению дефектами в ПО и при тестировании, анализ полученной информации и выявление общих причин и характерных шаблонов (паттернов), а так же плохих техник (анти-паттернов) описания требований, которые приводят к неполноте требований.

На данный момент путем анализа документации к нескольким промышленным системам выделены несколько признаков неполноты требований, например:

- При наличии условного элемента описаны не все ветви-последовательности событий.
- В одной ветви событий описывается изменение данных, но в параллельной ветви событий не описано, что происходит с этими данными.
- Упоминается функция, тип, объект, новое понятие, но нигде не описывается.

В результате исследования планируется составить некий контрольный список уточняющих вопросов (check-лист), получить набор паттернов и анти-паттернов, описывающих ситуации, в которых возможно проявление неполноты. Элементы этих списков не обязательно являются указателями на те места в требованиях, где требования неполны, они могут указывать на потенциально возможные моменты неполного описания свойств и функций системы. Использование таких списков и наборов шаблонов позволит найти уязвимые моменты в требованиях, пропущенные, не описанные, но возможные ситуации.

## Литература

1. В. В. Кулямин, Н. В. Пакулин, О. Л. Петренко, А. А. Сортов, А. В. Хорошилов, Формализация требований на практике, Препринт 13, ИСП РАН, Москва, 2006.

# Исследование и разработка методов повышения эффективности гибридных параллельных программ для высокопроизводительных систем

*Глазкова Екатерина Алексеевна*

*Студент*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,*

*Россия*

*email: [katrin-elinor@yandex.ru](mailto:katrin-elinor@yandex.ru)*

В настоящее время существует большой разрыв между пиковой производительностью суперкомпьютеров и той реальной производительностью, которую они показывают на тестовых пакетах и пользовательских задачах. Существуют различные подходы к созданию параллельных программ. Одними из наиболее распространенных подходов являются использование библиотеки MPI для систем с распределенной памятью и интерфейса OpenMP для систем с разделяемой памятью.

Современные компьютеры для высокопроизводительных вычислений, как правило, многоядерные, каждый узел в них представляет собой систему с разделенной памятью. В связи с этим кажется логичным использование гибридной MPI + OpenMP модели программирования на таких системах. В этой модели программирования MPI используется для организации взаимодействий между узлами (процессами), а OpenMP для многонитового программирования внутри узла.

Выбор между чистым MPI и гибридным MPI + OpenMP подходом при решении конкретной задачи является очень нетривиальным [2]. Следует учитывать влияние многих параметров, определяемых архитектурой машины, таких как стоимость коммуникаций, способ обращения к памяти и баланс производительности между процессором, памятью и коммуникационной средой.

Целью работы является сравнение и анализ эффективности выполнения набора модельных задач на высокопроизводительных системах с различной архитектурой и исследование и разработка методов повышения эффективности выполнения MPI и гибридных MPI + OpenMP программ на конкретной вычислительной системе на примере набора модельных задач.

Исследование проводится на системах Blue Gene/P и СКИФ-МГУ на трех тестовых задачах: стандартном бенчмарке (NAS Parallel Benchmarks), самостоятельно написанной тестовой программе (метод Якоби) и на программе, решающей задачу оптимизации управления развитием инфраструктуры типа поставщик-потребитель [1].

В работе проводится анализ и сравнение эффективности и масштабируемости набора тестовых задач на системах Blue Gene/P и СКИФ-МГУ, с целью проследить, как влияет архитектура системы на характер выполнения параллельной программы; сравниваются эффективность гибридной MPI + OpenMP модели программирования и чистой MPI-модели; исследуется влия-

ние настройки параметров задачи и параметров системы, на которой производилось тестирование, на эффективность гибридной модели программирования для различных тестовых задач.

### Литература

1. Жарков А.В. Пивоварчук Д.Г. Параллельный алгоритм решения задачи управления развитием инфраструктуры типа «поставщик-потребитель». // Программные системы и инструменты. Тематический сборник № 10. М.: Изд-во факультета ВМиК МГУ, 2009. С. 120-124.
2. Makris I. Mixed Mode Programming on Clustered SMP Systems. MSc in High Performance Computing. The University Of Edinburgh, 2005.

## О разметке области стека в задаче обратной инженерии

*Деревенец Егор Олегович*

*Студент*

*Московский государственный университет имени М.В. Ломоносова,*

*ВМК, Москва, Россия*

*email: yegord@ispras.ru*

При компиляции программ на языке высокого уровня локальные переменные и аргументы функции обычно отображаются в непрерывные участки памяти, расположенные по константным смещениям относительно регистра, указывающего на начало стекового кадра функции. Ставится задача восстановления данного отображения: по листингу ассемблерной программы требуется построить множество локальных переменных исходной программы и взаимно однозначное соответствие между переменными и соответствующими им участками памяти.

Идентификация локальных переменных и аргументов функций, имеющих тип массива или структуры, представляет определённую сложность: обращения к полям структур на стеке неотличимы от обращений к локальным переменным базовых или указательных типов, а информация о границах структур и массивов отсутствует в низкоуровневой программе. Методы решения данной задачи исследованы недостаточно.

Работа посвящена методу восстановления информации о расположенных на стеке локальных переменных структурных типов и массивов, реализованном в декомпиляторе TuDec, разрабатываемом в Институте системного программирования РАН.

Метод использует модель восстановления типов данных, описанную в [1]. Модель оперирует понятием объекта, соответствующим понятию локальной переменной в высокоуровневом языке. Метод решает задачу в несколько этапов.

На первом этапе выполняется итеративное построение множества объектов базовых и указательных типов и вычисление точных значений (абсолютных значений и смещений относительно указателя на начало стекового кадра функции), принимаемых каждым объектом.

На втором этапе для каждого объекта вычисляется надмножество принимаемых им значений, представляемое в виде двух интервалов: интервал абсолютных значений и интервал смещений относительно начала стекового кадра.

На третьем этапе строится карта стека. Она представляет собой разбиение стекового кадра на непересекающиеся области памяти. Области размечаются как соответствующие объектам структурных типов, массивам; определяются размеры массивов.

Использование предлагаемого метода в задаче декомпиляции позволяет увеличить количество восстанавливаемых локальных переменных и повысить качество восстановления программ.

### Литература

1. K. Dolgova, A. Chernov. Automatic Type Reconstruction in Disassembled C Programs. // Proceedings of the WCRE'08, October 2008.

## **Система справедливого планирования задач пользователя, на основе генетического алгоритма и унифицированного запуска задач через веб-интерфейс на разных вычислительных комплексах.**

**Князев Николай Александрович**

*Студент*

*Московский государственный университет имени М.В. Ломоносова,*

*ВМиК, Москва, Россия*

*email: irumata@gmail.com*

В данной работе рассматривается инструментальная система, обеспечивающая унифицированный запуск программ на суперкомпьютерах с различными Системами Пакетной Обработки (СПО) и справедливое разделение ресурсов между задачами пользователей, в предположении создания собственной высокоуровневой политики планирования по отношению к локальным СПО. Многие из современных подходов к планированию используют приоритеты задач[2]. В работе предлагается разработанный автором подход вычисления приоритета задач, заключающийся в создании невозобновляемых фишек приоритета[2].

Проблема планирования является NP-полной, для решения таких задач используют эвристические алгоритмы. В работе был использован генетический алгоритм, успешно применяемый при построении расписаний для неод-

народных систем[1]. Пусть на вход подаётся набор из  $n$  задач, с фишками приоритета  $p_i$  для каждой из них. Указывается количество процессоров  $s_i$  и максимальное время выполнения  $t_i$ . В качестве генотипа берётся расписание для  $n$  задач, представленное следующим образом:  $((\alpha_1, \tau_1), (\alpha_2, \tau_2), \dots, (\alpha_n, \tau_n))$ , где  $|\tau_i| \leq MAXVAL$  и  $|\alpha_i| \leq MAXVAL$ , - задают важность учёта числа процессоров и максимального времени исполнения задачи. Константа  $MAXVAL$  задаётся в параметрах алгоритма. Имея этот генотип расписание строиться с помощью алгоритма Backfill[2], при этом для каждой задачи  $i$  используется приоритет  $P_i$ , вычисляемый по формуле  $P_i = p_i \cdot t_i^{\tau_i} \cdot s_i^{\alpha_i}$ . Далее расписание, построенное по алгоритму Backfill, оценивается с помощью функции качества, задаваемой администратором системы, например: минимизировать время выполнения всего расписания. В конце происходит минимизация числа используемых фишек, с сохранением порядка запуска задач.

Представленный алгоритм интегрирован с Web-интерфейсом системы построения множественных выравниваний[3] Система состоит из Web-интерфейса, ядра на языке Python и bash-скриптов, запускающих задачу на суперкомпьютере, что обеспечивает независимость системы от количества суперкомпьютеров и используемых на них СПО(см Рис.). В текущий момент на уровне bash-скриптов поддерживаются такие СПО как: LoadLeveler, Torque, Cleo.

## Литература

- 1 Гаврилюк А. Б. Алексеев В.А. Метод оптимального статического планирования задач в распределенных вычислительных системах с использованием генетического алгоритма. // 2004 Киев, Проблеми програмування
- 2 Князев Н.А., Сальников А.Н. Управление динамическими приоритетами в планировании задач на многопроцессорных комплексах. // Программные системы и инструменты. Тематический сборник № 10. М.: ВМиК МГУ, 2009.
- 3 А.Н. Сальников «Интернет-сервис для построения множественного выравнивания последовательностей на многопроцессорных системах, созданный на основе data-flow модификации алгоритма MUSCLE» // Труды конференции ПАВТ-2009: ЮУрГУ, Челябинск

## Слова благодарности

Автор выражает благодарность научному руководителю Сальников Алексею Николаевичу.

## Иллюстрации



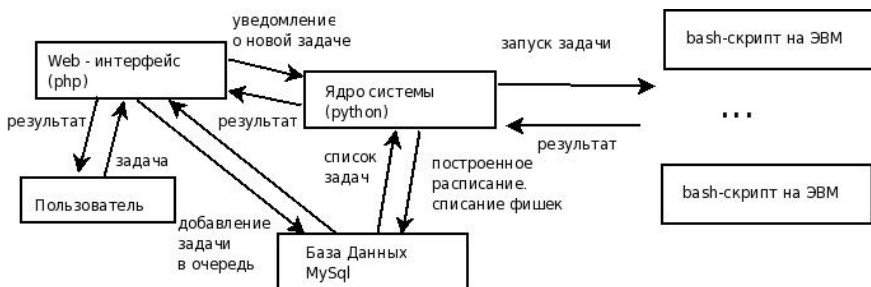


Рис. 1. Структура Система

## Библиотека Libjit для динамической компиляции

*Кононенко Кирилл Владимирович*

*Аспирант*

*Национальный Исследовательский Ядерный Университете*

*"МИФИ Факультет Кибернетики, Москва, Россия*

*email: Kirill.Kononenko@gmail.com*

Динамическая компиляция является основополагающим элементом технологий, обеспечивающих оптимизацию работы виртуальных машин и комплексных систем. Существуют три направления применения динамической компиляции. Во-первых, это оптимизация выполнения программы для повышения производительности. Во-вторых, это динамическая верификация поведения программ и исправление их поведения. В-третьих, это компиляция и оптимизация промежуточного представления программ для кластерных вычислений.

Целью было поставлено создание таких средств динамической компиляции [1 - 3], которые позволяют абстрагироваться от узкой специфики заданного приложения. Для этого был разработан набор стандартных средств компиляции, используемых во время выполнения программы. Он даёт достаточно большую степень свободы разработчикам, поскольку позволяет освободить их от малопродуктивной траты времени на низкоуровневую генерацию машинного кода. Использование предлагаемых средств также предоставляет возможность экспертам по созданию компиляторов времени выполнения и по безопасности сконцентрироваться на решении задач генерации и выполнения машинного кода, вместо высокоуровневого программирования.

Динамическая компиляция является основой для современных технологий программирования. Она необходима для реализации динамических и полудинамических языков. Более того, она позволяет эффективно выполнять байт-код виртуальной машины общего или специализированного назначения на регистровом процессоре. В случае доступности промежуточного представления программы, динамическая компиляция позволяет повысить производи-

тельность статически сгенерированного машинного кода. Возможность компиляции «на лету» существенно повысит и производительность приложений различного назначения.

Однако большинство существующих подходов к динамической компиляции испытывает ряд серьезных проблем: зависимость реализации от операционной системы, специфики виртуальной машины, библиотек операционной системы, сборщика мусора и, в целом, особенностей байт-кода. Наибольший объем работы компиляторов «на лету» связан с арифметическими операциями, с преобразованием типов данных, с доступом к памяти, с анализом графа управления и графа потока данных и генерацией корректного машинного кода. В результате большая часть работы над созданием компилятора времени выполнения программы, которая выполнена для одного приложения, не может быть использована для другого приложения.

В данном контексте была поставлена задача с помощью библиотеки Libjit создать средство, которое можно использовать для автоматизации генерации машинного кода при динамической компиляции и которое было бы лишено указанных (выше) ограничений.

### Литература

1. Документация libjit linear scan: <http://www.libjit.com>
2. К. В. Кононенко. Быстрая компиляция за время  $O(n)$  // Материалы Девятой международной конференции-семинара Высокопроизводительные параллельные вычисления на кластерных системах.
3. Документация libjit: [http://www.gnu.org/software/dotgnu/libjit-doc/libjit\\_1.html](http://www.gnu.org/software/dotgnu/libjit-doc/libjit_1.html)

### Слова благодарности

В разработке Libjit также приняли участие Rhys Weatherley, Norbert Bollow, Klaus Treichel, Aleksey Demakov, Avinash Atreya, Ankur Arora, Praveen Arimbrahodiyl, Christian Kolmeyer, Thomas Cort.

## Определение стратегии вытеснения Pseudo-LRU на ветвях бинарного дерева

*Корневич Евгений Валерьевич*

*Соискатель*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: kornevgen@gmail.com*

Ключевые слова: кэширование, кэш-память [2], стратегия вытеснения.

Актуальность исследования стратегии вытеснения Pseudo-LRU определяется ее применением в таких архитектурах как Pentium [1] и PowerPC [5].

Каноническое определение Pseudo-LRU предполагает оперирование с упорядоченным бинарным деревом высоты  $\log_2 w$  ( $w$  - (постоянный) оазмер кэш-памяти, оно должно быть степенью двойки). В результате обращения к элементу кэш-памяти изменяются некоторые метки вершин, на основе меток определяется путь к вытесняемому элементу [3].

Для целей исследования свойств отдельного элемента кэш-памяти (например, для генерации тестов[4]) было найдено другое определение стратегии вытеснения Pseudo-LRU. Каждому элементу кэш-памяти соответствует битовая строка длины  $\log_2 w$  (каждый бит соответствует вершине пути от корня того же упорядоченного бинарного дерева к этому элементу). Все элементы (листья) также можно пронумеровать битовыми строками длины  $\log_2 w$ . Определение Pseudo-LRU будет вестись для произвольного фиксированного элемента кэш-памяти, обозначим его номер (битовую строку) как  $\alpha$ . Далее каждое обращение в кэш-память будем кодировать номером (битовой строкой) искомого при этом элемента ( $\beta$ ). Это обращение будет изменять сопоставленную с элементом битовую строку. Тогда пусть  $\gamma$  - это значение этой строки для  $\alpha$  и осуществляется обращение к элементу с номером  $\beta$ . Разделим строку  $\alpha \oplus \beta$  на части - левая состоит из одних нулей (пусть ее длина  $n$ ) и средняя состоит только из одного бита - 1. Тогда в результате обращения новое значение строки  $\gamma$  составляется из трех частей - левая состоит из одних нулей и имеет длину  $n$ , средняя состоит из одной единицы, а остальные копируются из старого значения строки  $\gamma$  по тем же индексам. Вытесняемым  $\alpha$  будет в том случае, когда  $\gamma$  состоит из одних единиц. Это изменение можно выразить формулой  $\gamma' = ((\gamma \& (2 \cdot \delta - 1)) \mid \delta)$ , где  $\delta = 2^{\lceil \log_2 \alpha \oplus \beta \rceil}$ . Визуально можно представить себе изменение  $\gamma$  как перекрашивание вершин пути к  $\alpha$  - при обращении по пути к  $\beta$  перекрашиваем в белый вершины пути к  $\alpha$  до тех пор, пока они совпадают, и красим в черный вершину, на которой происходит расхождение путей.  $\alpha$  вытесняется, если весь путь к нему покрашен в черный.

### Литература

1. Таненбаум Э. Архитектура компьютера. Изд-во Питер, 2009.
2. Олифер В.Г., Олифер Н.А. Сетевые операционные системы. Изд-во Питер, 2008.
3. Tabak D. Advanced microprocessors. Mcgraw-Hill, 1995.
4. Dandamudi S.P. Fundamentals of computer organization and design. Springer, 2003.
5. Kornikhin E. SMT-based Test Program Generation for Cache-memory Testing // East and West-2009. С. 124-127.

# Решение задачи N тел на гибридных архитектурах с использованием библиотеки `ttgLib`

*Кривов М.А., Гризан С.А.*

*Студент*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,*

*Россия*

*email: m\_krivov@cs.msu.su*

В связи с широким распространением многоядерных процессоров и программируемых графических ускорителей производительность современного персонального компьютера стала достаточной для большинства задач, решаемых ранее исключительно с помощью суперкомпьютеров. Чтобы использовать все возможности современных компьютеров, были разработаны различные языки программирования и библиотеки, позволяющие создавать эффективные параллельные программы. К сожалению, большинство из них ориентируется на использование либо центрального процессора, либо графического ускорителя, что существенно усложняет создание приложения, использующего более одного вычислительного устройства.

В настоящей работе на примере решения задачи N тел рассматривается подход, позволяющий одновременно использовать центральный и графические процессоры. Данный подход реализован в библиотеке `ttgLib` [1] в виде специального параллельного примитива, интегрирующего технологии NVidia CUDA и Intel Threading Building Blocks. Основная идея заключается в представлении программы как ориентированного графа, каждому ребру которого сопоставляется поток данных, а каждой вершине — обработчик данных. Основным отличием данного подхода от других потоко-ориентированных моделей является возможность добавление для конкретных вершин дополнительных обработчиков, использующих нужную вычислительную технологию. Подобное представление позволяет не только распределять загрузку между всеми вычислителями, но и обеспечивать работоспособность в случае отсутствия программируемого графического ускорителя.

Для тестирования данной модели было рассмотрено два параллельных алгоритма решения задачи N тел. Первый из них был реализован с использованием библиотек Intel TBB, NVidia CUDA и разрабатываемой библиотеки `ttgLib`. Ускорение, полученное от использования только четырёхядерного центрального процессора (Intel TBB-версия), составило около 236%, что связано с увеличением числа операций по сравнению с последовательной версией. При совместном использовании центрального и графического процессоров (`ttgLib`-версия) полученное ускорение выросло до 796%. При использовании же только графического процессора (CUDA-версия) ускорение составило 1042%.

Согласно полученным результатам, использование гибридной архитектуры вида CPU + GPU понизило общую производительность примерно на 30% от достижимой, что объясняется накладными расходами из-за увеличения обращений к видеопамяти. При этом стоит отметить, что совместное

использование центрального и графического процессоров позволяет реализовать второй, более сложный рекурсивный алгоритм с меньшей алгоритмической сложностью, не имеющий очевидной реализации при использовании только графического ускорителя. После данной модификации ускорение ttgLib-версии составило 1120%.

### Литература

1. Кривов М.А., Припула М.Н., Конвейерная модель представления параллельных программ //Высокопроизводительные вычисления на кластерных системах — Материалы Девятой международной конференции-семинара, Владимир, 2009, С. 254-257.

## **Автоматическое динамическое распараллеливание вычислений для гетерогенных многопроцессорных компьютерных систем**

*Левченко Роман Игоревич*

*Аспирант*

*Киевский Национальный Университет имени Тараса Шевченко,*

*Радиофизический факультет, Киев, Украина*

*email: rmm01@mail.ru*

Сегодня развитие большинства направлений науки не возможно без использования многопроцессорных компьютерных систем. К примеру, современные исследования болезни Паркинсона[1] напрямую связаны с компьютерным моделированием. В результате развития многопроцессорных систем разработчики параллельных программ все чаще сталкиваются с проблемами гетерогенности, причина связана с расширением исследовательских кластеров. Для решения проблемы написания параллельных программ предлагается использовать авторское решение систему DDCI[2], которая сегодня активно развивается и уже используется в исследованиях болезни Паркинсона и сейчас планируется ее использование для экологического моделирования. Интегрирование DDCI системы в последовательную программу производится с помощью пометки распараллеливаемых функций алгоритма и добавления дополнительных библиотек. В предлагаемой системе построение графа распараллеливаемой части программы производится на одной или нескольких «серверных» вычислительных станциях. Серверные станции кроме построения графа являются ответственными за процесс распределения графа программы по всем «вычислительным» станциям и за балансировку вычислительной нагрузки на станциях. Вычислительные станции в системе используются только для выполнения запланированных расчетов и сами не реализуют никакой логики планирования, а вместо этого только пассивно отчитываются о своей загруженности серверным станциям. Рассматриваемая схема с разделением параллельной программы на серверные и вычислительные потоки показала свою эффективность на гетерогенных компьютерных системах. Так

размещая серверные потоки на самой высокопроизводительной станции можно добиться максимальной скорости анализа графа программы и как следствие максимальной расчетной нагрузки на все вычислительные станции. При этом за счет централизованной балансировки нагрузки система всегда будет в первую очередь выдавать расчетные задания для самых производительных вычислительных станций и использовать самые медленные только в крайних случаях, а так же проводить перераспределение вычислительной нагрузки в реальном времени. Такая политика распределения нагрузки на вычислительных станциях позволяет добиваться высокой эффективности от гетерогенных компьютерных систем. В заключении хочется сказать, что предложенный подход упрощает решение проблем распределенного хранения и маршрутизации данных.

### Литература

1. R.I. Levchenko, O.O. Sudakov, Yu.L. Maistrenko. Parallel software for modeling complex dynamics of large neuronal networks // Proc. 17th International Workshop on Nonlinear Dynamics of Electronic Systems, Rapperswil, Switzerland, June 21-24, 2009, P. 34-37.
2. R.I. Levchenko, O.O. Sudakov, S.D. Pogorelij. DDCI: Simple Dynamic Semiautomatic Parallelizing for Heterogeneous Multicomputer Systems // Proceedings of the 5-th IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications. Rende (Cosenza), Italy. September 21-23, 2009. pp. 208-211.

## Применение генетических алгоритмов в задаче проектирования отказоустойчивых вычислительных систем

*Пашков В.Н., Волканов Д.Ю.*

*Студент*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,*

*Россия*

*email: vasyapashkov@gmail.com*

В развитии современных вычислительных систем прослеживаются тенденции к повышению их функциональности за счёт увеличения числа используемых аппаратных модулей и компонентов, повышения сложности и размеров программного обеспечения, а также за счёт значительного роста данных, с которыми этим системам приходится работать. Проблема обеспечения стабильного безотказного функционирования вычислительных систем приобретает особую актуальность.

Один из наиболее часто применяемых на практике подходов к повышению надежности вычислительных систем основан на добавлении программных и аппаратных компонентов, дублирующих функциональность отдельных

частей системы, в сочетании с различными механизмами отказоустойчивости. Однако это приводит к повышению стоимости системы. Задача поиска конфигурации вычислительной системы с максимальной надежностью при заданном ограничении на стоимость является NP-трудной [1].

Для решения поставленной задачи использовалась классическая модель генетического алгоритма [2]. Критерием оценки качества решения (конфигурации системы) являлась целевая функция, равная надежности всей системы.

Целью данной работы является разработка и исследование эффективности гибридного генетического алгоритма, полученного путем модификации классической модели с учетом особенностей задачи.

В гибридном генетическом алгоритме используется схема нечеткой логики, которая после каждой итерации алгоритма анализирует изменение среднего значения целевой функции в популяции решений и изменение значения целевой функции лучшего решения. В результате чего происходит автоматическая перенастройка параметров алгоритма, изменение степени влияния операторов на генетический процесс.

Предложенный алгоритм тестировался на исходных данных, представленных в [2], и результаты его работы превзошли результаты работы классического генетического алгоритма по качеству получаемого решения.

### Литература

1. Chern M.S., "On the Computational Complexity of Reliability Redundancy Allocation in a Series System", Operation Research Letter, SE-13, pp. 582-592.
2. Wattanapongsakorn N. and Coit D.W., "Fault-tolerant embedded system design and optimization considering reliability estimation uncertainly", Reliability Engineering and System Safety, 92, pp. 395-407, 2007.

## Тестирование соответствия почтовых протоколов сети Интернет

*Тугаенко Анастасия Николаевна*

*Аспирант*

*Институт системного программирования РАН, Отдел Технологий  
программирования, Москва, Россия*

*email: tugaenko@ispras.ru*

В сети Интернет почтовое сообщение на пути от отправителя к получателю может проходить через цепочку промежуточных серверов, которые пересылают сообщение между различными зонами. Общая надежность механизма передачи электронной почты в значительной степени зависит от совместимости различных реализаций протоколов между собой. Проверка степени соответствия реализации требованиям стандарта – тестирование соответствия – служит одним из основных средств обеспечения совместимости.

В работе описывается метод тестирования соответствия, основанный на формальных спецификациях. Метод использует разработанную в Институте системного программирования РАН технологию автоматизации тестирования UniTESK [1], в которой требования стандарта формализуются посредством контрактных спецификаций, а тест задается как некоторый конечный автомат (автомат теста). Процесс тестирования представляет собой автоматический обход автомата теста, при котором наблюдаемое поведение реализации верифицируется на соответствие формальной спецификации. Использование формальных спецификаций позволяет автоматизировать как проверку правильности поведения реализации, так и оценку полноты тестирования, а представление теста как автомата дает возможность генерировать длинные и разнообразные последовательности тестовых событий. Разработанный метод содержит следующие основные шаги:

1. знакомство с предметной областью, разработка примеров и элементарных тестов;
2. анализ требований, построение упрощенной модели протокола; создание автомата теста с единственным состоянием;
3. выделение основных состояний протокола, построение автомата теста; добавление в тест блока, ответственного за перевод модельной системы из одного состояния в другое;
4. выделение и формализация требований, перенос проверки полученных от реализации откликов в отдельный модуль – спецификационный класс;
5. расширение автомата теста и спецификации для покрытия всех требований, формализованных в спецификации;
6. выполнение тестов и анализ полученных результатов.

С использованием инструмента JavaTESK[2], реализующего технологию UniTESK, был разработан тестовый набор для тестирования соответствия почтовых серверов стандартам протоколов SMTP и POP3. Разработанный тестовый набор применялся для тестирования нескольких open-source почтовых серверов. Для ряда реализаций были обнаружены несоответствия стандартам, включая:

- отсутствие поддержки обязательных команд;
- нарушения правил протокола (подача команд в недопустимых состояниях);
- неверные значения кодов откликов на команды протокола;
- заикливание при переадресации сообщений.



## **Литература**

1. Баранцев А.В., Бурдонов И.Б., Демаков А.В., Зеленов С.В., Косачев А.С., Кулямин В.В., Омельченко В.А., Пакулин Н.В., Петренко А.К., Хорошилов А.В. Подход UniTesK к разработке тестов: достижения и перспективы // Труды ИСП РАН, Т.5, 2004.
2. JavaTESK: первое знакомство. М: 2008. 28 с.

## **Слова благодарности**

Автор выражает признательность научному руководителю к.ф.-м.н. Пакулину Н.В. за помощь в проведении работы и подготовке тезисов.

# Дискретная математика и математическая кибернетика

## Легкотестируемые схемы для дизъюнкции

*Беджанова Светлана Руслановна*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,*

*Механико-математический факультет, Москва, Россия*

*email: azjunja@mail.ru*

Будем рассматривать схемы из функциональных элементов в произвольном полном конечном базисе, реализующие дизъюнкцию  $n$  переменных, где  $n$  не менее трех. В схемах допускается инверсная неисправность выхода ровно одного из элементов, то есть, если в исправном состоянии элемент реализует функцию  $h$ , то при его поломке на выходе элемента реализуется  $\bar{h}$ .

Пусть  $S$  — произвольная схема, реализующая функцию  $f(\vec{x}) = x_1 \vee \dots \vee x_n$ . Функция, реализуемая на выходе схемы при наличии в схеме неисправного элемента, называется функцией неисправности. Пусть  $g_1, \dots, g_k$  — все возможные попарно различные нетривиальные функции неисправности; функция неисправности называется нетривиальной, если она отлична от исходной функции, реализуемой исправной схемой. Множество  $T$  булевых наборов длины  $n$  называется единичным диагностическим тестом, если для любой пары функций из  $\{g_1, \dots, g_k, f\}$  в  $T$  найдется набор, на котором значения этих функций различны. Длиной теста называется число наборов в нем. Среди всех схем, реализующих  $f(\vec{x}) = x_1 \vee \dots \vee x_n$ , будем выделять те, которые допускают тесты минимально возможной длины. В данной работе при исследовании единичных тестов рассматриваются неизбыточные схемы.

**Теорема.** *Дизъюнкция  $n$  переменных может быть реализована неизбыточной схемой, допускающей единичный диагностический тест не более чем из двух наборов.*

Установлено, что оценка, приведенная в утверждении теоремы, в общем случае нелучшаема.

## Литература

1. Лупанов О.Б. Асимптотические оценки сложности управляющих систем. Изд-во МГУ. М., 1984.
2. Редькин Н.П. Дискретная математика. Изд-во ЦПИ при механико-математическом факультете МГУ. М., 2007.
3. Редькин Н.П. Надежность и диагностика схем. Изд-во МГУ. М., 1992.
4. Редькин Н.П. Единичные проверяющие тесты для схем при инверсных неисправностях элементов. Математические вопросы кибернетики. (2003) 12, 217-230.

### Слова благодарности

Автор выражает огромную благодарность за постановку задачи и внимание к работе своему научному руководителю профессору Редькину Николаю Петровичу.

## Свойства равновесий в коалиционных играх с неравномерным распределением агентов

*Вартанов Сергей Александрович*

*Студент*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: sergvart@gmail.com*

В работе рассматривается модель эндогенного формирования коалиций. Исследование проводится с помощью математического аппарата теории игр, в частности, понятий равновесия Нэша и коалиционного равновесия. В отличие от более ранней литературы ([1]) по данной тематике, рассматривается модель без побочных платежей, описываемая некооперативной игрой. Множество игроков предполагается достаточно большим, что позволяет рассматривать непрерывное распределение их по идеальным точкам. Фактически, исследуемая модель является обобщением модели с равномерным распределением, рассмотренной в [2], на случай произвольного непрерывного распределения.

Игроки характеризуются непрерывным распределением с функцией плотности  $f(x)$  по идеальным точкам в одномерном пространстве политических программ. Они выбирают стратегии из множества  $I^0 = \{0, 1, \dots, M\}$ , где  $M$  - достаточно большое целое число. Если выбрана стратегия  $i \neq 0$ , то считается, что игрок присоединился к коалиции  $i$ , иначе считается, что он воздержался от присоединения. В работе рассматриваются ситуации, в которых для каждой коалиции функция  $f_i(x)$  плотности распределения игроков, присоединившихся к ней, интегрируема. Итоговая политика  $p_i$  коалиции определяется как медиана этого распределения. Функция выигрыша  $U(x, V_i, p_i)$  игрока

возрастает по размеру  $V_i$  коалиции, к которой он присоединился, и убывает по расстоянию от итоговой политики  $p_i$  до его идеальной точки  $x$ .

В соответствии с определением, равновесием Нэша называется набор стратегий, в котором каждый игрок присоединяется к коалиции, в которой он имеет наибольший выигрыш при условии, что остальные игроки не изменяют свой выбор. Регулярным называется равновесие, в котором для любых различных коалиций значения итоговых политик различны. В работе исследуются вопросы существования, единственности и построения регулярного равновесия Нэша для функции выигрыша вида  $U(x, V_i, p_i) = R(V_i) - L(|x - p_i|)$ , где  $L(\cdot)$ ,  $R(\cdot)$  - строго возрастающие функции, а также вопрос устойчивости регулярных равновесий относительно раскола и локального объединения для  $R(r) = r$  и  $L(x) = a \cdot x + b \cdot x^2$ .

### Литература

1. Alberto Alesina, Enrico Spolaore. On the number and size of nations // The Quarterly Journal of Economics, Vol. 112, No. 4. (Nov., 1997), pp. 1027-1056.
2. Alexander Vasin, Denis Stepanov. Endogenous formation of political parties // Mathematical and Computer Modelling 48 (2008) pp. 1519-1526

### Слова благодарности

Автор благодарит профессора, д.ф.м.н. Васина А.А. за помощь в подготовке тезисов.

## О семействах точных вершинных $k$ -расширений графов при $k > 1$

*Долгов Александр Алексеевич*

*Аспирант*

*Саратовский государственный университет имени Н.Г.*

*Чернышевского, Факультет компьютерных наук и информационных*

*технологий, Саратов, Россия*

*email: Dolgov.A.A@gmail.com*

Граф  $H$  называется *точным вершинным  $k$ -расширением* (ТВ- $k$ Р) графа  $G$ , если граф  $G$  изоморфен каждому подграфу  $H$ , получающемуся путем удаления любых его  $k$  вершин и всех связанных с ними дуг (ребер). В данном докладе будут рассмотрены семейства точных вершинных  $k$ -расширений при  $k > 1$ .

Первые два бесконечные семейства ТВ- $k$ Р графов при любом  $k > 0$  были предложены Харари и Хейзом в статье 1996 года [3], это семейство вполне несвязных графов  $O_n$  и семейство полных неориентированных графов  $K_n$ .

Позже М.Б. Абросимовым было показано, что среди неориентированных графов не существует других графов, кроме  $O_n$  и  $K_n$ , имеющих ТВ- $k$ Р при  $k$

$> 1$  [1]. Также удалось установить связь между ориентированными и неориентированными ТВ-кР, которая заключается в том, что симметризация ТВ-кР орграфа, всегда является ТВ-кР уже неориентированного графа [2].

В ходе исследований, удалось доказать следующий факт:

**Теорема.** Орграф  $G = (V, \alpha)$  является ТВ-кР тогда и только тогда, когда части графа  $G^* = (V, \alpha^*)$ , где

$$\alpha^* = \{(x, y) | x, y \in V, (x, y) \in \alpha \& (y, x) \in \alpha\}$$

и  $G^{**} = (V, \alpha^{**})$ , где

$$\alpha^{**} = \{(x, y) | x, y \in V, (x, y) \in \alpha \& (y, x) \notin \alpha\}$$

также являются ТВ-кР.

Это утверждение означает, что если в заданном графе, который является ТВ-кР, выделить симметричную часть, образованную из пар встречных дуг, и направленную часть, образованную из дуг, не имеющих встречных, то каждая из этих частей также должна быть ТВ-кР.

Очевидно, что никакой граф с  $n$  вершинами не может содержать в себе часть изоморфную  $K_n$ , кроме самого графа  $K_n$ . Это означает, что среди орграфов имеющих встречные дуги не может быть графов, имеющих ТВ-кР при  $k > 1$ .

Среди диграфов М.Б. Абросимовым было выделено семейство ТВ-кР при любом  $k > 0$ , это семейство транзитивных турниров  $T_n$  [2].

Удалось показать, что  $T_n$  – это единственное семейство ТВ-кР, все представители которого являются бесконтурными графами. Любое другое ТВ-кР обязательно будет сильносвязным графом. Также удалось показать, что среди сильносвязных турниров не может существовать графа, являющегося ТВ-кР при любом  $k > 1$ .

Таким образом, существует только три семейства точных вершинных  $k$ -расширений графов при любом  $k > 0$ . Это полные и вполне несвязные графы в неориентированном случае и транзитивные турниры для случая диграфов.

### Литература

1. Абросимов М.Б. Минимальные расширения дополнений графов // Теоретические задачи информатики и ее приложений. Саратов:СГУ. 2001. Вып.4. С.11-19.
2. Абросимов М.Б. Минимальные расширения транзитивных турниров // Вестник ТГУ. Приложение. 2006. № 17. С. 187 – 190.
3. Harary F., Hayes J.P. Node fault tolerance in graphs // Networks. 1996. V. 27. P. 19-23.

# Репрезентативность исходов аукциона функций предложений с учетом неопределенности спроса

*Долматова Марина Станиславовна*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,*

*Россия*

*email: ms.marina.dolmatova@gmail.com*

Рынки электроэнергии и газа, как правило, организованы в форме аукциона единой цены, где заявка производителя задается монотонной функцией, определяющей зависимость продаваемого объема электроэнергии (газа) от цены. Важной особенностью таких рынков является неопределенность спроса, связанная со случайными колебаниями из-за изменения внешних условий и значительным изменением объема потребления в течение суток. Это необходимо учитывать при моделировании аукциона.

В работе мы рассматриваем аукцион функций предложения. Равновесием в функциях предложения является набор заявок, каждая из которых обеспечивает максимизацию прибыли участника при фиксированных заявках соперников при любом значении спроса (см. [2]).

Можно ли ожидать, что исходы, получаемые в данной модели, соответствуют реальному поведению игроков на рынке? Ответ мы получаем, исследуя адаптивный процесс (динамику наилучших ответов, ДНО) на предмет сходимости набора стратегий к SFE для двух вариантов рынков (см. [1]). Рассматриваем симметричную дуополию с линейной функцией спроса и предельными издержками: А) линейными, Б) постоянными при ограничении производственной мощности  $Q$  (см. [3]).

В работе показано, что в случае А на каждой итерации существует заявка наилучшего ответа и выведен ее вид. Доказано, что ДНО сходится к SFE при отсутствии ограничений на значение спроса. При анализе ДНО для случая Б оказывается существенным соотношение максимального значения  $D_{\max}$ , принимаемого спросом, и значения  $Q$ : если ограничение мощности существенно, то ДНО сходится к SFE, равному равновесию Вальраса, если несущественно, то нет. Для случая Б, наиболее типичного для реальных рынков электроэнергии, реальное поведение игроков не будет соответствовать SFE.

Исходя из полученных результатов, рассмотренная модель аукциона и равновесие в функциях предложения не являются репрезентативными для анализа и предсказания исходов на реальных рынках электроэнергии и газа.

## Литература

1. Васин А.А., Васина П.А. Модели конкуренции функций предложения и их приложение к сетевым аукционам // EERC, Москва, 2005.
2. Klemperer P., Meyer M. Supply function equilibria in oligopoly under uncertainty // *Econometrica*, 1989, 57 (6), 1243-1277.

**Получение «справедливых» устойчивых дележей в  
задачах кооперативных игр**

**Клименко Оксана Александровна**

*Аспирант*

*Московский педагогический государственный университет,  
математический, Москва, Россия*

*email: akasana@nm.ru*

Кооперативные игры являются разделом теории игр, в котором основным вопросом является не выбор оптимальных стратегий, а установление разумного определения выигрышей игроков, вступающих в коалиции.

Целью представленной в докладе работы является рассмотрение возможности коррекции исходной характеристической функции для получения устойчивого дележа, который, кроме того, обладал бы еще и свойством «справедливости». Устойчивыми являются дележи  $C$ -ядра, а относительной «справедливостью» обладают вектор Шепли и  $N$ -ядро. Проблема заключается в том, что для некоторых игр  $C$ -ядро пусто, вектор Шепли, в общем случае, не принадлежит  $C$ -ядру, а  $N$ -ядро, являющееся селектором  $C$ -ядра, существует и в играх с пустым  $C$ -ядром.

Одним из путей нахождения устойчивого «справедливого» дележа является получение непустого  $C$ -ядра в игре, где в качестве решения может быть выбрано  $N$ -ядро. Для игры, заданной в форме характеристической функции, получаем следующую задачу коррекции по минимуму евклидовой нормы:

$$\tilde{v}(S) = v(S) + \Delta_{v(S)} \quad \sum_{S \subset N} \Delta_{v(S)}^2 \rightarrow \min$$

Аналогично рассматривается решение задачи для минимаксного критерия.

$$\tilde{v}(S) = v(S) + \Delta_{v(S)} \quad \max_S (\Delta_{v(S)}) \rightarrow \min, \quad \forall S \subset N$$

Предлагается в качестве альтернативного решения рассмотреть  $C_p$ -ядро:

$$C_p = \left\{ x \mid \|\Delta_S(x)\|_p \rightarrow \min, \quad \forall S \subset N, \quad \sum_{i=1}^n x_i = v(N) \right\},$$

где  $\|\cdot\|_p$  – норма Гельдера,  $\Delta_S(x) = \max(0, e(S, x))$ ,  $e(S, x) = v(S) - \sum_{i \in S} x_i$ .

Другой вариант получения устойчивого и «справедливого» дележа – перемещение вектора Шепли в  $C$ -ядро. Рассматривая такую возможность, с учетом коррекции данных с теми же критериями, была решена задача нахождения расстояния между ними.

Проведен анализ влияния коррекции на положение коалиций для различных игровых ситуаций, в том числе для игр, в которых есть носитель. Рассмотрена интерпретация полученных результатов.

### Слова благодарности

Автор выражает благодарность д.ф.-м.н., профессору Горелику В.А. за научное руководство и постановку задач.

## Инъективная $L(2, 1)$ -раскраска расщепляемых неразложимых униграфов

*Максимович Олег Валерьевич*

*Аспирант*

*Белорусский Государственный Университет,*

*Механико-математический факультет, Минск, Белоруссия*

*email: aleh.maksimovich@gmail.com*

Впервые задача о  $\lambda$ -раскраске появилась, по-видимому, в [5] и [7]. Сегодня уже имеется ряд глубоких результатов, связанных с  $\lambda$ -раскрасками. В частности, доказана NP-полнота задачи распознавания " $\lambda(G) \leq c$ " при фиксированном  $c \geq 4$  и выделены некоторые специальные классы графов, для которых эта задача остается NP-полной ([4], [5]). Среди них оказался популярный класс расщепляемых графов [2] (для которого, как известно, классическая задача "раскраска" решается за время  $O(n)$ ).

В [3] представлен алгоритм, строящий оптимальную раскраску в классе пороговых графов за время  $O(n)$ . Там же приведен эвристический алгоритм, решающий ту же задачу в классе матрогенных графов. В [1] для задачи "инъективная раскраска" приводится новая адекватная интерпретация как оптимизационной задачи на множестве перестановок вершин графа.

Предлагаемая работа – продолжение [1]. Это первая часть работы " $L(2, 1)$ -раскраска униграфов", где для класса униграфов решаются задачи  $\lambda$ -раскраски и инъективной  $\lambda$ -раскраски ( $\lambda'$ -раскраски).

Вкратце, наша идеология в следующем. Опираясь на характеристику униграфов и теорию декомпозиции из [6], разложить произвольный униграф на неразложимые части, затем для каждой из этих частей найти оптимальную  $\lambda'$ -раскраску, "склеить" эти раскраски в оптимальную  $\lambda'$ -раскраску исходного униграфа и, наконец, переработать последнюю в оптимальную  $\lambda$ -раскраску. В соответствии с этой идеологией здесь строится оптимальная  $\lambda'$ -раскраска лишь для тех униграфов, которые являются базовыми графами неразложимых операторов, т.е. для нетривиальных неразложимых униграфов.

Базой для нашего метода служит следующее:

1. Операторная декомпозиция графов,
2. Упомянутая выше новая интерпретация задачи  $\lambda'$ -раскраски из [1].

### Литература



1. Максимович О. В., Тышкевич Р. И. Инъективная  $L(2, 1)$ -раскраска как оптимизационная задача на множестве перестановок вершин графа: доминантно-пороговые графы // Труды Института математики НАН Беларуси. 2009. Том 17. № 1. С. 110-118.
2. Bodlaender H.L., Kloks T., Tan R.B., van Leeuwen J.  $\lambda$ -coloring of graphs. // Tech Report. Dept. of Computer Science, Utrecht University, 1999.
3. Calamoneri T., Petrechi R.  $\lambda$ -Coloring matrogenic graphs // Discrete Applied Mathematics. 2006. V. 154. P. 2445–2457.
4. Fiala J., Kloks T., Kratochvil J. Fixed-Parameter Complexity of  $\lambda$ -Labelings // Lecture Notes in Computer Science. 1999. V. 1665. P. 350–363.
5. Griggs J.R., Yeh R.K. Labelling graphs with a condition at distance 2 // SIAM Journal of Discrete Mathematics. 1992. V. 5(4). P. 586–595.
6. Tyshkevich R. I. Decomposition of graphical sequences and unigraphs // Discrete Mathematics. V. 220. 2000. P. 201–238.
7. Yeh R.K. Labelling graphs with a condition at distance two. Ph.D. Thesis. Department of Mathematics, University of South Carolina, 1990.

## Алгоритмы решения систем булевых уравнений с использованием ассоциативных принципов обработки информации.

*Мелузов Антон Сергеевич*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: [asmelouzov@mail.ru](mailto:asmelouzov@mail.ru)*

В настоящее время проблема решения систем алгебраических уравнений над конечными полями исследуется достаточно широко. Эта задача является NP-трудной и полиномиальный алгоритм её решения в настоящее время не известен. Однако, в процессе криптоанализа регулярно возникает необходимость решения таких систем с фиксированной структурой (число и вид уравнений, число переменных и др.).

В работе рассмотрены системы булевых уравнений вида  $\{f_1(X_1) = 0, f_2(X_2) = 0, \dots, f_m(X_m) = 0\}$ , где  $X_i \subseteq X = \{x_1, \dots, x_n\}$ , и исследован алгоритм решения, близкий к использованным в работах [2], [3], но предназначенный для реализации на ассоциативных вычислителях, описанных, например, в [1]. Пусть каждая из функций  $f_i$  выбрана случайно равновероятно из всего множества булевых функций от переменных  $X_1$ . Пусть  $d_1 = |X_1|, d_2 =$

$|X_2 \setminus X_1|, d_3 = |X_3 \setminus (X_1 \cup X_2)|, \dots, d_m = |X_m \setminus (\bigcup_{i=1}^{m-1} X_i)|, S_i = \sum_{j=1}^i d_j$ . На полуинтервале  $(0, m]$  определим функции  $d(x) = d_{\lceil x \rceil}, S(x) = \int_0^x d(y) dy, P(x) = S(x) - x$ .

В работе выделены классы систем уравнений указанного вида, для которых доказаны следующие субэкспоненциальные асимптотические оценки математического ожидания трудоёмкости алгоритма поиска решений.

**Теорема 1** *Рассмотрим системы уравнений в которых, выполнены следующие условия: последовательность значений  $d_1, d_2, d_3, \dots, d_k, \dots, d_m$  — невозрастающая,  $d_1 > 1, m \geq n, m = O(n)$  при  $n \rightarrow \infty$ , а функция  $d(x): d(x) \leq C \cdot \sqrt{\frac{n}{\log n}} \cdot \frac{1}{\sqrt{x}} = \delta(x), x \in (0, k_0]$ , где  $k_0$  такое, что  $d_{k_0} > 1$  и  $\forall k > k_0 \Rightarrow d_k \leq 1$ , а  $C$  — некоторая константа.*

*Тогда математическое ожидание трудоёмкости алгоритма поиска всех решений таких систем уравнений имеет верхнюю асимптотическую оценку  $O(n \cdot 2^{C^2 \cdot \frac{n}{\log n}})$  при  $n \rightarrow \infty$ .*

**Теорема 2** *Если в условиях Теоремы 1 функция  $d(x) : d(x) \leq C \cdot \frac{\sqrt{n}}{x+1} = \delta(x), x \in (0, k_0]$ , а  $C$  — некоторая константа, то математическое ожидание трудоёмкости алгоритма поиска всех решений таких систем уравнений имеет верхнюю асимптотическую оценку  $O(n \cdot 2^{C\sqrt{n}(\log C + \log n - 1)})$  при  $n \rightarrow \infty$ .*

## Литература

1. Фостер К. Ассоциативные параллельные процессоры. М., Энергоиздат, 1981.
2. Semaev Igor. On solving sparse algebraic equations over finite fields I // Proceedings of WCC07, Versailles, France, INRIA, 2007. С. 361-370.
3. Semaev Igor. On solving sparse algebraic equations over finite fields II: <http://eprint.iacr.org/2007/280>.

## Расшифровка обобщенных псевдо-булевских монотонных функций

**Осокин Виктор Владимирович**

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,*

*Механико-математический факультет, Москва, Россия*

*email: [osvic@mail.ru](mailto:osvic@mail.ru)*

В работе рассмотрена задача параметро-эффективной расшифровки псевдо-булевских функций в ее стандартной постановке [1,2]. Введен класс ГРМ псевдо-булевских функций, такой что если функция из этого класса принимает одно и то же значение на двух сравнимых наборах булевого куба, то она принимает то же значение и на всей грани, образованной этими

двумя наборами. Класс GPM содержит, в частности, класс псевдо-булевских монотонных функций, класс разбивающих функций SPL [2] и, следовательно, класс булевских монотонных функций M, класс симметричных монотонных (пороговых) функций THR и класс дизъюнкций переменных OR. Получен алгоритм параметро-эффективной расшифровки GPM. Доказаны нижние оценки, показывающие оптимальность этого алгоритма на упомянутых классах.

Пусть  $\Phi$  — некоторый класс псевдо-булевских функций. Через  $\Phi^n$  обозначим подмножество этого класса, состоящее из функций от не более чем  $n$  переменных, а через  $\Phi^{k,n}$  — подмножество  $\Phi^n$ , функции которого имеют не более  $k$  существенных переменных. Пусть  $\varphi(A, f)$  — число запросов на значение функции, которое требуется алгоритму  $A$  для расшифровки функции  $f$ ,  $\mathcal{A}(\Phi)$  — множество алгоритмов расшифровки класса  $\Phi$ . Положим  $\varphi(\Phi, n, k) = \min_{A \in \mathcal{A}(\Phi)} \max_{f \in \Phi^{k,n}} \varphi(A, f)$ . Величину  $\max_{f \in \Psi} \varphi(A, f)$  назовем сложностью алгоритма  $A$  на конечном классе  $\Psi$ .

**Лемма 1.** Пусть  $\Phi \subseteq \text{GPM}$ . Пусть имеется алгоритм  $A \in \mathcal{A}(\Phi)$ , сложность которого на  $\Phi^n$  не превосходит  $\psi(n)$  по порядку при  $n \rightarrow \infty$ . Тогда существует алгоритм расшифровки  $A' \in \mathcal{A}(\Phi)$ , сложность которого на  $\Phi^{k,n}$  не превосходит  $k \log n + \sum_{i=1}^k \psi(i)$  по порядку при  $n, k \rightarrow \infty$ .

**Лемма 2.** Имеет место  $\varphi(\text{OR}, n, k) \gtrsim k \log \frac{n}{k}$ ,  $\varphi(\text{M}, n, k) \gtrsim k \log \frac{n}{k} + \sqrt{\frac{2}{\pi} \frac{2^k}{\sqrt{k}}}$ ,  $\varphi(\text{SPL}, n, k) \gtrsim \max(k \log(n - k + 1), 2^k)$  при  $n, k \rightarrow \infty$ .

**Теорема.** Имеет место  $\varphi(\text{GPM}, n, k) \asymp \varphi(\text{PM}, n, k) \asymp \varphi(\text{SPL}, n, k) \asymp k \log n + 2^k$  при  $n, k \rightarrow \infty$ ,  $\varphi(\text{M}, n, k) \asymp k \log n + \frac{2^k}{\sqrt{k}}$  при  $n, k \rightarrow \infty$ ,  $\varphi(\text{THR}, n, k) \asymp \varphi(\text{OR}, n, k) \asymp k \log n$  при  $n \rightarrow \infty$  и  $\log k = o(\log n)$ .

## Литература

1. Кудрявцев В.Б., Гасанов Э.Э., Подколзин А.С. Введение в теорию интеллектуальных систем. М., 2006.
2. Осокин В.В. О параллельной расшифровке разбиений булевого куба // Интеллектуальные системы. Т.13. 2009.

## Слова благодарности

Автор выражает благодарность профессору Э.Э.Гасанову за постановку задачи и помощь в работе.

# Равновесия двухэтапного рынка со случайным исходом на спотовом рынке.

*Шарикова Агата Андреевна*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: agatha.sharikova@ubs.com*

Форвардный рынок – известный механизм снижения рыночной власти компаний. Это важный факт, так как для электроэнергии характерна высокая концентрация производства, а потребители часто не имеют рыночной власти. Интересным является результат, полученный Бушнеллом [1]. Он рассмотрел двухкратный аукцион Курно в условиях линейных предельных издержек и отсутствия арбитража. Автор показал, что возможность заключения форвардных контрактов снижает рыночную власть производителей так же, как увеличение числа производителей на рынке с  $n$  до  $n^2$ . Существуют проблемы, связанные с исследованием [1]: фактическая динамика цен не соответствует гипотезе о равенстве цен на спотовом и форвардном рынках, а приоритет богатых потребителей при покупке товара на форвардном рынке не имеет места на реальных рынках.

В данной работе исследуется двухэтапная модель олигополии со случайной ценой на спотовом рынке. На рынке действуют риск-нейтральные арбитражеры, цена  $p^f$  на форвардном рынке равна математическому ожиданию цены  $p^s$  на спотовом рынке. Анализируется стратегическая модель взаимодействия производителей, арбитражеров и потребителей. На первом шаге каждая фирма  $a$  назначает объем  $q_a^f$  предложения на форвардном рынке. На втором шаге каждый потребитель  $b$  принимает решение о покупке товара на форвардном рынке, исходя из его функции полезности, характеризующейся резервной ценой  $r_b$  и параметром  $\lambda_b$ , показывающим степень избегания риска потребителем. Потребители, решившие покупать товар на форвардном рынке, формируют функцию спроса  $D^f(p)$ . Цена  $p^f$  определяется из соотношения  $D^f(p^f) = \sum_a q_a^f$ . На третьем шаге происходит выбор объемов  $q_a^s$  предложения на спотовом рынке. Происходит аукцион Курно, определяется цена  $p^s$ :  $D^s(p^s) = \sum_a q_a^s$ . На спотовом рынке с вероятностью  $w$  реализуется цена  $p1$ ,  $p1 \leq p^f$ , с вероятностью  $1 - w$  - цена  $p2$ ,  $p2 \geq p^f$ . Оптимальные стратегии производителей ищутся как совершенные подыгровые равновесия модели. Доля риск-предпочитающих агентов с  $r_b > p2$  считается постоянной.

Модель позволила охарактеризовать равновесное поведение потребителей. На форвардном рынке товар покупают только риск-избегающие потребители с  $r_b > p^f$ . В равновесии производители используют коррелированные смешанные стратегии. Наблюдается понижение цен и рост объема выпуска.

## Литература

1. Bushnell J. Oligopoly equilibria in electricity contract markets. CSEM

Working Paper, University of California Energy Institute, 2005.

# Дифференциальные уравнения, динамические системы и оптимальное управление

О задаче Коши для нелинейного уравнения  
соболевского типа с переменным коэффициентом  
*Аристов Анатолий Игоревич*

*Соискатель*

*Московский государственный университет имени М. В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: ai\_aristov@mail.ru*

Работа посвящена исследованию асимптотики при больших временах решения следующей задачи Коши:

$$\frac{\partial}{\partial t} (\Delta u - u) + a (\Delta u - u) + \mu(x) |u|^\sigma u = 0, \quad u(x, 0) = u_0(x) \quad (1)$$

Здесь  $u(x, t)$  – действительная функция,  $x \in R^N$ ,  $t \geq 0$ . Будем предполагать, что  $\mu(x) \in L_q(R^N)$ , причем если  $N \leq 3$ , то  $2 \leq q \leq \infty$ , если  $N \geq 4$ , то  $N/2 < q \leq \infty$ . Пусть, кроме того,  $a > 0$ ,  $\sigma \geq 1$ ,  $u_0(x) \in L_\infty(R^N) \cap H^2(R^N)$ .

Уравнение описывает нестационарные процессы в полупроводниках. Получают развитие идеи из [1], где изучалось уравнение  $\frac{\partial}{\partial t} (\Delta u - u) + a \Delta u + |u|^\sigma u = 0$ . Показано, что если начальные данные достаточно малы в смысле некоторой нормы, то качественное поведение решения при  $t \rightarrow \infty$  определяется линейными членами уравнения.

Формально проведем следующие преобразования. Положим  $u = v e^{-at}$ , применим к уравнению оператор  $B[\cdot] = F^{-1} \left[ \frac{1}{1+|p|^2} F[\cdot] \right]$  ( $F$  – преобразование Фурье) и проинтегрируем от 0 до  $t$ .

Получим следующее интегральное уравнение:  $v(x, t) = u_0(x) + \int_0^t e^{-\sigma a \tau} B[\mu |v|^\sigma v] d\tau$ . Введем пространство  $X = C[0, \infty; L_\infty(R^N) \cap H^2(R^N)]$  с нормой

$\|\varphi\|_X = \sup_{t \in [0, \infty)} (\|\varphi\|_{L^\infty(\mathbb{R}^N)} + \|\varphi\|_{H^2(\mathbb{R}^N)})$ . Будем называть обобщенным решением данной задачи Коши такое  $u \in X$ , что  $u = ve^{-at}$ , где  $v \in X$  – решение названного интегрального уравнения.

**Теорема.** *Существует такая постоянная  $K > 0$ , что если  $\|u_0\|_{L^\infty(\mathbb{R}^N)} + \|u_0\|_{H^2(\mathbb{R}^N)} \leq K$ , то обобщенное решение задачи существует и единственно, причем при  $t \rightarrow \infty$  имеет место следующее асимптотическое представление:  $u(x, t) = A(x)e^{-at} + O(e^{-(\sigma+1)at})$  (функция  $A$  зависит только от пространственной переменной).*

### Литература

1. N. Hayashi, E. Kaikina, P. Naumkin, I. Shishmarev. Asymptotics for Dissipative Nonlinear Equations. Springer-Verlag, 2006.
2. G. N. Watson. A treatise on the theory of Bessel functions. Cambridge University Press, Cambridge, England, 1944.

## Вероятностная модель межрегиональных отношений в контексте поставок природного газа

*Борисова Анастасия Викторовна*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: asik.vmk@gmail.com*

Работа посвящена актуальным проблемам моделирования энергетической безопасности. В ней построена и исследуется математическая вероятностная модель межрегиональных отношений и связанных с ними поставок природного газа от поставщика ( $S$ ) к транспортеру ( $T$ ) и затем к конечному потребителю ( $C$ ). Она отражает влияние политических взаимоотношений стран, входящих в состав цепочки поставки, на надежность поставки энергоносителей.

В основу исследования положена стилизованная модель, опирающаяся на предположение марковости перехода [1,2] в следующее состояние в дискретном времени. В качестве фазовых переменных модели используется вектор  $r^k = (r_{st}^k, r_{tc}^k, r_{sc}^k)$ , каждая компонента которого отражает попарные политические взаимоотношения между двумя из трех стран в момент времени  $k$ .  $r_{ij}^k$  принимает значение 0, в случае негативных отношений, или 1, в случае позитивных отношений. Переменная  $g^k$  в бинарном смысле отражает наличие поставки газа от поставщика к потребителю. Определены условные вероятности, которые в стилизованном виде формализуют практику принятия политических решений участниками процесса.

Цель работы состоит в исследовании вероятностных характеристик перехода трехкомпонентной системы в состояние, которое условно назовём коллапсом. Коллапс – состояние системы, при котором прекращается поставка газа. В работе предложены два дополняющих друг друга подхода к описанию поведения системы. В первом подходе с помощью исходно заданных условных вероятностей строится матрица переходных вероятностей  $A = [a_{ij}]_{i,j=1}^8$  и вектор-столбец вероятностей перехода в состояние коллапса  $q = [q_i]_{i=1}^8$ . Выведены явные формулы для вероятности наступления коллапса на шаге  $k$  и до шага  $k$ , получены их оценки сверху. Проведено исследование асимптотического поведения системы в предположении положительности матрицы  $A$  и доказано, что вероятность наступления коллапса возрастает и стремится к 1 на бесконечности.

Второй подход состоит в рассмотрении пространства расширенных состояний системы, описывающих одновременно и политические отношения, и наличие поставки газа. В нём строится матрица переходных вероятностей  $\tilde{B}$  специального вида, который помогает отделить состояние наступления коллапса от неколлапсных состояний. Получены новые представления формул вероятностей наступления коллапса, формула для математического ожидания номера шага, на котором наступает коллапс.

В работе исследована возможность упорядочивания состояний системы с точки зрения вероятности коллапса на последующих шагах, проведено многочисленное количество экспериментов с помощью компьютерного моделирования.

Методы и подходы, предложенные в данной работе, допускают интерпретацию в терминах других приложений марковских процессов и могут найти применение в различных задачах, не обязательно связанных с поставкой газа.

### Литература

1. А.С.Выдрин, А.В.Михалёв. Стохастические матрицы и анализ защищенности автоматизированных систем //Фундаментальная и прикладная математика, 13, 2007.
2. Ф.Р.Гантмахер. Теория матриц. М.: Наука, 1966.

### Оценивание рисков в экологических сетях

*Вещинская Виктория Валерьевна*

*Студент (специалист)*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет Вычислительной Математики и Кибернетики, Москва,  
Россия*

*email: mail@gmail.com*

Одной из глобальных экологических проблем является сокращение биологического разнообразия. Последствиями исчезновения биологических видов становится не только потеря уникального генетического материала, но и



дестабилизация экосистем, угроза другим биологическим видам. Важный вопрос в экологии – оценка влияния вымирания отдельных видов, участвующих в сложных пищевых цепочках, на другие виды и на функционирование системы в целом. Возможны ситуации, когда исчезновение некоторой популяции приводит к изменению диеты остальных участников экосистемы и переходу всей системы в новое стабильное состояние, либо к ее полной деградации.

Математически, данная задача относится к оцениванию неравновесной динамики в связанных графах [1]. Основными переменными в задаче являются стоки энергии, запасённые в узлах графа, и потоки энергии между ними. Система является открытой, то есть существуют энергетические связи между её узлами и окружающей средой. Стартовый режим функционирования системы – равновесный, то есть стоки и потоки энергии не меняются со временем. Система функционирует в дискретном времени, перемещение порции энергии (потока) от одного узла к другому занимает один временной интервал. Далее предполагается вымирание одного узла системы в начальный момент времени и рассматривается динамика системы в сделанном предположении. Такая динамика называется коллапсной.

В данной работе рассматриваются два метода оценивания ущерба. Первый из них, метод дисконтированного оценивания ущерба для экосистем от деградации того или иного её участника, предложен в работе [3]. Он основан на сравнении запасов энергии при равновесной и коллапсной динамике функционирования системы. Второй – метод сетевого анализа экосистем, описан в работе [2] и заключается в изучении полных энергетических связей между узлами системы.

В работе приводятся результаты моделирования динамики экосистем, проводится указанными выше методами оценка влияния деградации отдельного вида пищевой сети на остальных членов сообщества и на функционирование системы, исследуется вопрос устойчивости равновесного состояния экосистем, рассматриваются различные подходы моделирования энергетических потоков в экосистемах.

Модели рассматриваемых экосистем предоставлены Программой Evolution and Ecology Program, International Institute for Applied Systems Analysis, Laxenburg, Austria.

### Литература

1. Оре О. Теория графов. М. Наука, 1980.
2. Patten B. C. Energy, emery and environs // Ecol. Modell. 1992. V.62, p. 29-69.
3. Rovenskaya E., Kryazhimskiy A. Collapse Assessment for Energy Networks (draft of a paper). 2009.
4. Tilly L. J. The structure and dynamics of Cone spring // Ecological Monographs. 1968. V. 38:2, p. 169-197.

# Аппроксимация множеств достижимости управляемых систем в среде Matlab

Винников Евгений Владимирович

Аспирант

Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия

email: [evinnikov@gmail.com](mailto:evinnikov@gmail.com)

При анализе поведения управляемых систем важную роль играют множества достижимости (МД). Рассматривается нелинейная управляемая динамическая система, описываемая уравнением  $\dot{x} = f(t, x, u)$ ,  $u \in P$ ,  $t \in I$ ,  $I = [0, T]$ ,  $x(0) \in X_0$ ,  $x(t) \in F$ . (1)

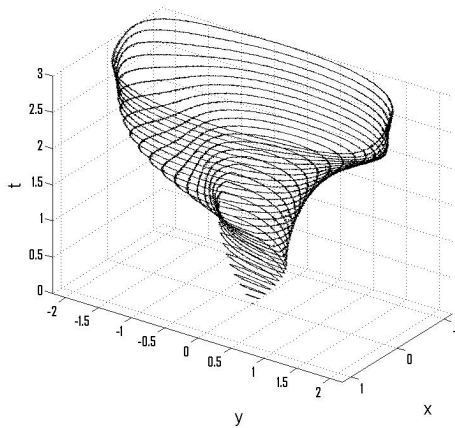


Рис. 1. Интегральная воронка

Универсальным численным методом аппроксимации МД нелинейных управляемых систем является метод пикселей [1, 2, 4]. Данный метод позволяет строить МД, множества управляемости и интегральные воронки для нелинейных управляемых систем, описываемых системой дифференциальных уравнений с управляемыми параметрами, как с фазовыми ограничениями, так и без них. Для МД, построенных методом пикселей, были получены оценки расстояния Хаусдорфа между точным МД и его дискретной аппроксимацией.

В среде *Matlab* разработана программа *PixelSet* для построения МД и интегральных воронок для задачи (1) с фазовыми ограничениями и ограничениями на управление в виде неравенств, содержащая библиотеку из более 50 задач. Разработанная программа позволяет строить МД двухмерных и

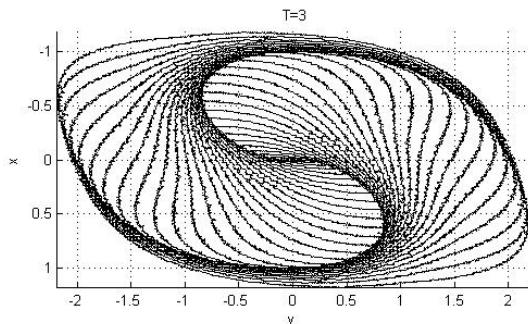


Рис. 2. Проекция интегральной воронки на плоскость  $Oxy$

трёхмерных управляемых систем. В трёхмерном случае реализована триангуляция поверхности; имеется возможность построения сечения поверхности произвольной плоскостью, что позволяет получить более полную информацию о характере поведения трёхмерного МД.

В качестве примера работы программы *Pixelset* построена интегральная воронка нелинейной управляемой системы, описываемой уравнением Дуффинга [3]

$$\begin{cases} \dot{x} = y, & x(0) = 0, \quad |u| \leq 1, \\ \dot{y} = -x - x^3 + u, & y(0) = 0, \quad 0 \leq t \leq T, \end{cases}$$

моделирующим поведение нелинейной жесткой пружины, при  $T=3$  (рис. 1, 2).

### Литература

1. Винников Е. В. Численное построение множества достижимости нелинейных управляемых систем. // Обзорные прикладной и промышленной математики. 2009. Т. 16. Вып. 5. С. 824–825.
2. Гусейнов Х.Г., Моисеев А.Н., Ушаков В.Н. Об аппроксимации областей достижимости управляемых систем. // Прикладная математика и механика. 1998. Т. 62. Вып. 2. С. 179–187.
3. Ли Э.Б., Маркус Л. Основы теории оптимального управления. М., 1972. 574 с.
4. Никольский М.С. Об аппроксимации множества достижимости для

## Применение операторной обратной связи для экспоненциальной стабилизации билинейных систем

*Гончаров Олег Игоревич*

*аспирант*

*МГУ, ВМиК, Россия, Москва*

*email: goncharovo i@yandex. ru*

Рассматривается задача экспоненциальной стабилизации билинейной системы вида

$$\dot{x} = Ax + \sum_{i=0}^m u_i B_i x,$$

где  $x \in \mathbb{R}^n$  — известный вектор состояния системы, а  $u_1, \dots, u_m$  — управления. В случае устойчивой матрицы  $A$  задача имеет тривиальное решение в виде управления  $u = 0$ .

В литературе для стабилизации билинейных систем рассматривались возможности использования различных законов управления. Однако, линейные ( $y = cx$ , см. [2]) и квадратичные законы управления ( $y = x^T Q x$ , см. [3]) в общем случае не могут обеспечить сходимость системы к нулю, если матрица  $A$  неустойчива. Более того, квадратичный закон управления не гарантирует экспоненциальную устойчивость. Для использования равномерных законов управления (законы вида  $y = \frac{x^T Q x}{\|x\|^2}$ ) часто требуется нейтральная устойчивость матрицы  $A$  (например, [4]).

В данной работе рассматривается возможность использования динамической операторной (в терминах монографии [1]) обратной связи для сведения задачи экспоненциальной стабилизации исходной системы в нуле к задаче стабилизации преобразованной системы

$$\dot{z} = (A + \alpha I)z + \sum_{i=0}^m u_i B_i z$$

в ограниченной области диссипативности. Последняя задача может решаться, например, с использованием квадратичных законов управления.

### Литература

1. Емельянов С.В., Коровин С.К. Новые типы обратной связи. Управление при неопределенности. М., 1997.
2. Калман Р., Фабл П., Арбиб М. Очерки по математической теории систем. М., 1971.
3. Molher R.R. Nonlinear Systems. Vol 2. Application to Bilinear Control. New York and London, 1991.

4. Rahn C.D. Stabilization Conditions for Strictly Bilinear Systems with Purely Imaginary Spectra // IEEE Tr. on AC. 1998. Vol. 41. pp. 1346-1347.

## Метод имитационного моделирования групповой динамики толпы

*Гребенников Роман Владимирович*

*Аспирант*

*Воронежский государственный университет, Факультет*

*компьютерных наук, Воронеж, Россия*

*email: grebennikov.roman@gmail.com*

Большие скопления людей есть обычное явление в современном мире, и их моделирование является важным вопросом в различных областях, таких как компьютерная графика и анализ архитектурных объектов на предмет оптимальности людского потока и эвакуации в экстремальных ситуациях.

Предложенный в данной работе метод моделирования поведения толпы основан на мультиагентном подходе к непрерывному представлению набора локальных полей, каждое из которых характеризует один из факторов воздействия на участника толпы, а также его индивидуальные характеристики: поле глобальной цели движения, максимальной скорости, поле дискомфорта и т.д. Из данного набора полей для каждого участника толпы формируется потенциал  $\Phi()$ , при минимизации которого формируется вектор дальнейшего движения. Простейшим методом минимизации потенциала  $\Phi(x)$  является метод градиентного спуска:

$$\vec{x} = -f(x, \theta) \cdot \frac{\nabla\phi(x)}{\|\nabla\phi(x)\|}, \quad \|\nabla\phi(x)\| = C, \quad \Phi(x) = \int_P C ds, \quad C = \frac{\alpha f + \beta + \gamma g}{f} \quad (2)$$

где  $\vec{x}$  – вектор дальнейшего движения,  $\alpha, \beta, \gamma$  – коэффициенты,  $f(x, \theta), g(x)$  – используемые потенциальные поля максимальной скорости и дискомфорта,  $P$  – путь,  $ds$  – интеграл по длине пути. Тестирование данного метода на практике показало его свойство к попаданию в локальный минимум потенциального пространства. Для решения этой проблемы в данной работе предложено использовать метод оптимизации роя частиц [1], для поиска минимума потенциала  $\Phi()$  в области видимости каждого агента толпы, что позволит эффективно избегать попадания в локальные минимумы.

На основании предложенного в данной работе метода создана система моделирования поведения толпы *Cogetcrowd*, позволяющая как производить моделирование толпы в заданных условиях, так и сравнивать различные методы и подходы к моделированию. Данная система позволила произвести сравнение предложенного в данной работе метода с моделями, предложенными Treuille [3], и Reynolds [2], которое показало его преимущество по всем тестируемым параметрам: длина, кривизна пройденного пути и количество попаданий в локальные минимумы.

## Литература

1. Eberhart, R., Kennedy, J. A new optimizer using particle swarm theory // Proceedings of the 6th Int. Symposium on Micro Machine Science, Nagoya, Japan, 1995, p. 39-43.
2. Reynolds, C. Flocks, Herds, and Schools: A Distributed Behavioral Model // Computer Graphics, SIGGRAPH '87 Conference Proceedings, New York, Vol. 21, p. 25-34.
3. Treuille A., Cooper S., Popović Z. Continuum crowds // ACM SIGGRAPH 2006 Papers, New York, p. 1160-1168.

## Синтез оптимального управления в математической модели терапии вирусных инфекций

*Зайчик Светлана Юрьевна*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: szaychik@t48.ru*

В работе ставится и решается задача синтеза оптимального управления в математической модели, описывающей развитие под терапевтическим воздействием легко мутирующего вируса.

Пусть  $u_1(t)$  — количество единиц основного вируса,  $u_2(t)$  — количество единиц мутантного вируса,  $h(t)$  — количество используемого лекарственного средства.

Рассмотрим следующую систему:

$$\begin{cases} \dot{u}_1 &= \lambda_1 - \gamma_1 u_1 - \alpha_1 u_1 f_1(h), & u_1(0) = u_1^0; \\ \dot{u}_2 &= \lambda_2 + \alpha_3 f_3(h) - \gamma_2 u_2 - \alpha_2 u_2 f_2(h), & u_2(0) = u_2^0; \\ \dot{h} &= -\gamma_3 + W, & h(0) = h^0. \end{cases} \quad (3)$$

Здесь  $\lambda_1, \lambda_2$  — скорости воспроизводства основного и мутантного вирусов,  $\gamma_1, \gamma_2$  — коэффициенты смертности вирусов,  $\gamma_3$  — коэффициент диссипации терапевтического средства,  $f_1(h)$  и  $f_2(h)$  функции терапии, характеризующие интенсивность воздействия лекарственного средства,  $f_3(h)$  — функция, описывающая увеличение скорости воспроизводства мутантного вируса под действием лекарственного средства,  $W(t)$  — функция управления, задающая количество лекарственного средства, которое может быть введено в пациента в единицу времени,  $0 \leq W(t) \leq R$ ,  $R > 0$ ,  $\alpha_i, i = 1, 2, 3$  — положительные постоянные.

Ставится задача синтеза оптимального управления — управления с обратной связью. Для любого текущего состояния системы, описываемого равенствами  $t = \bar{t}$ ,  $u_1 = \bar{u}_1$ ,  $u_2 = \bar{u}_2$ ,  $h = \bar{h}$ , найти такое управление  $W^*(t, u_1, u_2, h)$ , удовлетворяющее условию

$$0 \leq W^*(t) \leq R,$$

при котором функционал

$$\Phi(t) = u_1^2 + \epsilon u_2^2 \quad (4)$$

достигал бы своей нижней грани в заданный момент времени  $t = T$ .

Для решения поставленной задачи применяется метод динамического программирования. При построении гладкой функции цены метод характеристик не даёт желаемого результата. В работе вводится понятие псевдорешений уравнения Гамильтона-Якоби-Беллмана. Такие решения не определяют функцию цены во всём расширенном пространстве. Данная проблема решается с помощью сингулярной характеристики и поиска особого оптимального управления. Интересным и неожиданным фактом является то, что построенное таким образом решение уравнение Гамильтона-Якоби-Беллмана оказывается везде гладкой функцией.

### Литература

1. Deforche K. Estimating the Relative Contribution of dNTP Pool Imbalance and APOBEC3G/3F Editing to HIV Evolution In Vivo / Koen Deforche, Ricardo Camacho, Kristel Van Laethem, Beth Shapiro, Yves Moreau, Andrew Rambaut, Anne-Mieke Vandamme, Philippe Lemey // Journal of Computational Biology. - 2007. - 14(8). - P. 1105-1114.
2. Субботин А.И. Обобщенные решения уравнений в частных производных первого порядка: перспективы динамической оптимизации. - Москва-Ижевск: Институт компьютерных исследований, 2003.
3. Братусь А.С., Антипов А.В. Оптимальная стратегия химиотерапии в математической модели динамики численности клеток неоднородной опухоли. // Вычислительная математика и математическая физика. — Том 49, №1, 2009.
4. Melikyan A.A. Generalized Characteristics of First Order PDEs: Applications in Optimal Control
5. and Differential Games. - Boston: Birkhauser, 1998.
6. Габасов Р., Кириллова Ф.М. Особые оптимальные управления. -Москва: Наука, 1973.

## Верхние оценки порядка стабилизатора для ММО систем

**Капалин Иван Владимирович**  
*Аспирант*

*Московский государственный университет имени М.В. Ломоносова  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: [ikapalin@gmail.com](mailto:ikapalin@gmail.com)*

Одной из важных задач в теории автоматического управления является

задача о получении верхних оценок порядка стабилизатора  $k$  для многомерных (ММО) систем вида:

$$\begin{cases} \dot{x} = Ax + Bu, \\ y = Cx, \end{cases} \quad (5)$$

где  $x \in \mathbb{R}^n$  - фазовый вектор системы,  $y \in \mathbb{R}^l$  - измеряемый выход,  $u(t) \in \mathbb{R}^m$  - управление,  $A, B$  и  $C$  - постоянные матрицы соответствующих размерностей,  $\text{rank } C = l, \text{rank } B = m$ . Важность таких оценок в прикладном применении, так как порядок стабилизатора определяет сложность схем его реализующих.

Самой простой верхней оценкой по праву можно считать

$$k \leq n - l,$$

которая связана с наблюдателем Люенбергера. Представляет интерес очевидный вопрос "можно ли улучшить эту оценку?". Ответ на этот вопрос неоднозначен. С одной стороны существуют системы, для которых нельзя построить стабилизатор ниже порядка Люенбергера  $n - l$ , с другой стороны существуют системы, для которых порядок стабилизатора может быть значительно понижен.

На основе работ [1] и [2] было показано, что следующие оценки порядка стабилизатора для систем вида (5) верны

$$k \leq n + 1 - l - m$$

$$k \leq \max \{n + 2 - l - m - \nu, \nu - l, \mu - m + 1\}$$

$$k \leq \max \{n + 2 - l - m - \mu, \mu - m, \nu - l + 1\},$$

где  $\nu$  и  $\mu$  - индекс управляемости и наблюдаемости системы (5), соответственно. Целью же этой работы является простое исследование связи между данными оценками и при возможности выделения наилучшей или наихудшей из них.

В результате исследования удалось показать, что ни одна из оценок не содержится в двух других. Оказалось, что оценка

$$k \leq n + 1 - m - l$$

лучше других в единственном случае, когда  $\nu = n - m + 1$  и  $\mu = n - l + 1$ . Были получены при фиксированных  $n, m$  и  $l$  области, в которых одна оценка превосходит другие две. Так же в работе показана область значений переменных  $\mu$  и  $\nu$ , для которых существуют системы вида (5). Приведен алгоритм построения систем для каждой пары  $(\mu, \nu)$  из этой области.

## Литература

1. Н. Kimura, "A further result on the problem of pole assignment by output feedback," IEEE Trans. Automatic Control, Vol.AC-22, pp.458-463, 1977.
2. Н. Kimura, "On pole assignment by output feedback," International J. Control, Vol.28, pp.11-22, 1978.



# Динамическое программирование в линейных системах с состояниями в виде распределений

*Мазуренко Станислав Сергеевич*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова*

*Факультет вычислительной математики и кибернетики, Москва,*

*Россия*

*email: stasmazurenko@mail.ru*

В теории оптимального управления важную роль играют детерминированные задачи оптимизации, в которых состояние системы точно известно в каждый момент времени. Динамика таких объектов описывается системой дифференциальных уравнений при заданном начальном положении. При добавлении неопределённостей в такие системы, появляются известные схемы интегрирования, наблюдения и управления [1].

Для проблем, возникающих в стохастической постановке, неопределённость моделируется при помощи вероятностных распределений для начальных состояний и входов. Весьма общая постановка задачи управления потоками моделируется в терминах уравнения Лиувилля [3], решение которого может пониматься в обобщённом смысле [2].

В данной работе метод динамического программирования применяется в анализе детерминированных задач управления для обыкновенных дифференциальных уравнений с состояниями в виде распределений. Выводятся аналогии уравнений Гамильтона-Якоби-Беллмана в общем случае, используя которые, разрешается задача оптимального управления линейными системами с квадратичным функционалом. Описание потоков, порождённых начальными распределениями, сосредоточенными на компактном множестве, приведены в работе [4].

## Литература

1. Куржанский А.Б. Управление и наблюдение в условиях неопределённости. М.: Наука, 1977.
2. Гельфанд И.М., Шилов Г.Е. Обобщенные функции и действия над ними. М.: Государственное издательство физико-математической литературы, 1959.
3. Brockett R.W. Optimal control of the Liouville equation. Providence, R.I.; American Mathematical Society; 1999.
4. Kurzbanki A.B., Filippova T.F. On the theory of trajectory tubes - a mathematical formalism for uncertain dynamics, viability and control. Advances in Nonlinear Dynamics and Control. Birkhauser, Boston, 1993, pp.122-188.

## Слова благодарности

Автор выражает признательность своему научному руководителю, Александру Борисовичу Куржанскому за постановку задачи и руководство в процессе её решения.

## **Оптимизация биотехнологического процесса с использованием методов математического моделирования**

*Плотникова Юлия Валентиновна*

*Аспирант*

*Тамбовский государственный университет имени Г.Р. Державина,  
Институт математики физики и информатики, Тамбов, Россия  
email: plotnikovay@yandex.ru*

Рассматриваемый биотехнологический процесс представляет собой комплексную технологию утилизации отхода производства этанола с выработкой на этой основе бактериальной биомассы.

Данный технологический процесс включает в себя несколько различных стадий: стадия получения бактериальной биомассы [1], стадия разделения ее в термофлотаторе [2], стадия гранулирования [3].

На этой основе созданы математические модели основных технологических единиц процесса – биотехнологического реактора, термофлотатора, сушилки-гранулятора. Данные модели адекватны реальным объектам и позволяют рассмотреть динамику изменения основных параметров технологических единиц (концентрации биомассы и субстрата, температуру, влажность готового продукта и т.д.). Применение численных методов позволяют получить результат оптимизации на основе актуальной экономической информации. Целью данной работы является решение задачи экономической оптимизации указанного технологического процесса с учетом современных цен на материалы, сырье, энергетические ресурсы, а также существующим уровнем оплаты труда персонала.

При решении задач оптимального проектирования и оптимизации биотехнологического в качестве целевой функции используется совокупность капитальных и эксплуатационных затрат. Использование такой целевой функции приводит к снижению себестоимости продукции, тем самым обеспечивая ее конкурентную способность на рынке, повышает экономию природных, энергетических и иных ресурсов, снижает затраты на монтаж и эксплуатацию оборудования. При решении задач оптимизации будем стремиться к минимизации данной целевой функции.

### **Литература**

1. Арзамасцев А.А. Оптимальное проектирование технологического процесса и аппарата для выработки бактериальной биомассы из отходов производства этилового спирта. – Тамбов, 1984.
2. Дудаков В.П. Математическое моделирование и оптимизация процесса термофлотационного разделения суспензий. – Тамбов, 2001.

3. Иньков В. И. Математическое моделирование и оптимизация статических режимов процесса грануляции в псевдооживленном слое. – Тамбов 1984.

## **О внутренних и внешних оценках информационных множеств задачи гарантированного оценивания**

*Прозоров Илья Сергеевич*

*Студент*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: ilya.s.prozorov@gmail.com*

Одним из наиболее востребованных направлений современной математики является теория динамических систем с неопределенностью. Особое внимание уделяется проблемам оценивания в задачах управления и наблюдения. В приложениях статистическое описание обычно не является полным и адекватным. По этой причине, в рамках теории гарантированного оценивания развивается альтернативный подход, связанный с предположением об ограниченности неопределенных параметров модели множествами известного вида.

В рамках направления гарантированного оценивания удается получить точные аналитические описания множеств возможных состояний изучаемой системы. Однако, как правило, эти множества имеют достаточно сложную структуру. Поэтому интерес вызывает задача построения для них гарантированных аппроксимаций в некотором классе множеств. В качестве таковых в работе используются эллипсоиды.

В настоящей работе рассматривается задача поиска внутренних и внешних оценок информационного множества по результатам доступных наблюдений, подверженных помехам. Основное внимание уделяется случаю вырожденной матрицы преобразования в уравнении наблюдений. В этом случае информационное множество является пересечением цилиндрических множеств. Схема решения строится на идее приближения этих цилиндрических множеств эллипсоидами. Внешние и внутренние оценки пересечения эллипсоидов построены в [1, 2, 3]. Обсуждаются вычислительные схемы. Приведены примеры численного моделирования.

### **Литература**

1. **Kurzhanski A. B., Vályi I.** Ellipsoidal calculus for estimation and control. SCFA. Boston: Birkhäuser, 1997.
2. **Kurzhanski A. B., Varaiya P.** On Ellipsoidal Techniques for Reachability Analysis. Part I: External Approximations // Optimization methods and software. 2002 V. 17. N. 2. P. 177-206.

3. Kurzhanski A. B., Varaiya P. On Ellipsoidal Techniques for Reachability Analysis. Part II: Internal Approximations, Box-Valued Constraints // Optimization methods and software. 2002 V. 17. N. 2. P. 207-237.

## Программное приложение для параметрической идентификации динамических моделей

*Семёнов Михаил Игоревич*

*Аспирант*

*Чувашский государственный университет имени И.Н. Ульянова,  
Радиотехнический факультет, Чебоксары, Россия  
email: arktos.mike@gmail.com*

В разработанном программном приложении применён и адаптирован под задачу параметрической идентификации алгоритм Левенберга-Марквардта в модификации Флетчера [3], оптимальный для решения переопределённых нелинейных систем. Задачи, решаемые данной системой, возникают при настройке и «доводке» различных моделей на экспериментальные или статистические данные [1]. Область применения данной программы довольно широка – от научных исследований с применением математического моделирования до использования при разработке систем автоматизации технологических процессов в промышленности.

Основной пакет MATLAB предоставляет средства для целевой оптимизации, например, функция `fminsearch` служит для нахождения локального минимума нелинейных функций. Однако в случае, когда задача задана переопределённой нелинейной системой уравнений, применение функции `fminsearch` является неэффективным, а функции `lsqnonlin`, входящей в библиотеку Optimization Toolbox - недостаточно гибким.

Реализация метода Левенберга-Марквардта в модификации Флетчера состоит из следующих стадий. Во-первых, иницируются дополнительные параметры по умолчанию или определенными пользователем значениями. Во-вторых, начальные условия для итерационного процесса определяются как сумма квадратов невязки, соответствующей начальному приближению решения и матрице Якоби в этой точке. Затем, наконец, запускается итерационный цикл выполнения алгоритма Левенберга-Марквардта в модификации Флетчера [2].

Разработанная система способна достаточно быстро и удобно осуществить подбор параметров заданной математической модели, при которых она наиболее точно описывает данные. Разработанный алгоритм выбора рабочего отрезка времени позволяет наиболее эффективно получить желаемые результаты. Программа имеет необходимые гибкие настройки и достаточно мощный аппарат предварительной обработки данных. Это значительно ускоряет процесс параметрической идентификации сложных систем. Система является унифицированной для применения в различных прикладных задачах науки и промышленности.

## Литература

1. Гилл Ф., Мюррей У., Райт М. Практическая оптимизация. М.: Мир, 1985.
2. Balda M. LMFsolve: Levenberg-Marquardt-Fletcher's algorithm for nonlinear least squares problem// MathWorks, MATLAB Central, File Exchange, Id=1606. 2007.
3. Fletcher R. A Modified Marquardt Subroutine for Nonlinear Least Squares// Rpt. AERE-R 6799, Harwell. 1971.

# Компьютерная графика

## Анализ изображений на основе характеристик краев и формы объектов

*Жариков Александр Леонидович*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,*

*Физический факультет, Москва, Россия*

*email: alexandr-j@yandex.ru*

Сравнение и классификация изображений может быть осуществлена с тем или иным успехом множеством методов. Как известно, методы морфологического анализа изображений [1] позволяют получить описание изображения, инвариантное к оптическим условиям его регистрации. Однако инвариантность к геометрическим преобразованиям объекта на изображении такому описанию изначально не свойственна, и для ее достижения требуются дополнительные приемы. В данной работе исследуются возможности использования для этой цели методов, основанных на анализе краев объектов на изображении. Цель работы – оценить возможности такого подхода и наметить пути его дальнейшего улучшения.

Структура краев, полученных тем или иным детектором границ в большинстве случаев достаточно сложна: края имеют ветвления и самопересечения. Однако для отдельного сегмента края, представимого параметрически заданной непрерывной кривой, легко построить дескриптор, инвариантный к таким геометрическим преобразованиям объекта, как сдвиг, плоский поворот и масштабирование. В данной работе такой дескриптор вычисляется на основе кривизны края. Для описания изображения такими дескрипторами был разработан метод выделения сегментов из произвольной структуры краев.

Таким образом, изображение описывается набором одномерных дескрипторов. Однако такое описание применительно к реальным фотографиям имеет ряд недостатков: во-первых, дескрипторов получается слишком много, причем большинство из них не информативно. Для решения этой проблемы в работе были исследованы возможности анализа информативности дескрипторов и их фильтрации. Во-вторых, края объектов могут иметь самые разнообразные формы. Это делает бессмысленным построение дескриптора на

основе точного определения кривизны, т.к. важную роль играет пространственный масштаб ее изменений: нужно избегать как слишком подробной, так и слишком малой детализации. Для разрешения этой ситуации на основе эмпирических данных была разработана схема построения дескрипторов, оптимальная для большинства интересных случаев. Наконец, метод игнорирует многие важные характеристики границ объектов (например, такие информативные элементы, как прямые линии и окружности, дают константные дескрипторы). В этой связи остается много возможностей для синтеза с другими методами анализа краев.

Полученное описание изображений может применяться для детектирования объектов на фотографиях, задач классификации или поиска копий изображений. В работе оцениваются возможности описанного подхода как в отдельности, так и в сочетании с методами морфологического анализа.

Алгоритмы были реализованы на языке Java. Эффективность алгоритмов оценивалась на небольшой, специально подготовленной базе изображений.

### Литература

1. Пытьев Ю.П., Чуличков А.И. Морфологические методы анализа изображений. М.:ФИЗМАТЛИТ, 2009.
2. Форсайт Д., Понс Ж. Компьютерное зрение. Современный подход. Вильямс., 2004.

## Система трёхмерной реконструкции объектов и калибровки проекторов с использованием структурированного света

*Зейналов Рамиз Шакирович*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,*

*Россия*

*email: ramiz.zeynalov@gmail.com*

Набирающие популярность системы расширенной реальности часто строятся на базе проекторов, проецирующих изображения или видео на объект. Ключевой задачей в данной области является построение трёхмерной модели объекта и определение взаимного расположения проектора и объекта. Для создания трёхмерных моделей объектов реального мира уже существуют разнообразные подходы, но они обладают рядом недостатков: либо требуют длительной работы, либо не дают удовлетворительных результатов, либо слишком дороги. Кроме того, существующие методы обычно не работают с динамическими объектами и сценами.

В качестве альтернативы были разработаны методики активного стереозрения, где кроме камер используются проекторы, проецирующие на объект изображение или последовательность изображений - шаблон структури-

рованного света, от которого зависят ограничения на моделируемые объекты, скорость получения модели и её точность [1]. При этом камера снимает изображения объекта с шаблоном. Это позволяет автоматически получить соответствия между точками шаблона и изображения с камеры и восстановить трёхмерную модель. Задача состоит в том, чтобы с помощью проектора и камеры восстановить трёхмерную модель объекта, рассчитать внутренние параметры проектора и камеры и их взаимное расположение и сопоставить полученную модель с уже имеющейся или ранее полученной.

Процесс решения задачи разбивается на этапы:

1. Внутренняя калибровка камеры. Для автоматического определения внутренних параметров камеры необходимо отснять порядка 20 фотографий калибровочного шаблона в виде шахматной доски.
2. Засветка объекта интереса шаблонами. После засветки каждым шаблоном делается снимок камерой.
3. Определение точечных соответствий. Используется зависящий от шаблона алгоритм его распознавания и построения соответствий.
4. Калибровка стереопары. После определения некоторого числа точечных соответствий между оконными системами координат стереопары используется классический алгоритм стереозрения Bundle Adjustment [2].
5. Построение плотной трёхмерной модели. После получения полных калибровок камеры и проектора и определения точечных соответствий между их оконными системами координат и калибровки системы трёхмерная модель может быть получена путём триангуляции [1].
6. Сопоставление модели с реальной сценой. Полученная модель сопоставляется с известной моделью объекта, то есть вычисляется такое преобразование модели (и калибровок проектора и камеры соответственно), что она максимально точно впишется в полученную модель объекта реального мира. Таким образом, в результате получают соответствия систем координат полученной и известной ранее моделей, что необходимо для задач проекционного дизайна и систем расширенной реальности.

### Литература

1. M.Mazaheri, M.Momeni “3D Modeling Using Structured Light Pattern and Photogrammetric Epipolar Geometry”, ISPRS, Beijing, 2008
2. Bill Triggs, Philip McLauchlan, Richard Hartley, Andrew Fitzgibbon “Bundle Adjustment — A Modern Synthesis”, LNCS, pp. 298-375. Springer Verlag, 2000



# Реконструкция текстур сильно загороженных фасадов зданий с использованием нескольких изображений.

*Кононов В.А., Коношин В.С.*

*Студент*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: kovlan@rambler.ru*

В настоящее время активно развивается область трехмерной реконструкции городов. Ее важной составной частью является создание реалистичных текстур фасадов зданий по фотографиям. Но на реальных снимках фасад часто загорожен различными объектами переднего плана, такими как деревья, провода. Для качественной реконструкции необходимо очищение текстуры фасадов. Современные методы очистки текстур работают медленно и не подходят для реального применения, так как они работают с любыми изображениями, не используя знания об объектах на них.

В случае зданий естественно использовать знание структуры фасада, которую проще всего задавать координатами и типами окон. Существуют несколько полуавтоматических и автоматических алгоритмов поиска такой информации [3].

Наши исследования показали, что при большом количестве объектов переднего плана, когда нет ни одного чистого окна, невозможно быстро и качественно восстановить текстуру здания. Поэтому в таких случаях необходимо воспользоваться дополнительной информацией. Наиболее естественным дополнением являются фотографии отдельных окон вблизи. Причем достаточно сфотографировать по одному окну каждого типа.

Предлагаемый метод работает в два этапа. Вначале происходит развертка фотографий отдельных окон с целью приведения горизонтальных и вертикальных линий фасада к горизонтальным и вертикальным линиям на изображении. Для этого используется существующий алгоритм автоматической развертки [1]. Второй шаг заключается в нахождении масштаба и сдвига для вклейки отдельных фотографий на основную текстуру. Алгоритм пробегается по дискретному множеству масштабов и сдвигов. Для каждой пары параметров вырезаются области всех окон фасада, которые накладываются на изображение отдельного окна, и проводится сравнение. Было протестировано несколько метрик для такого сравнения. Лучший результат показало сравнение направлений градиентов. Для корректного подсчета близостей карт направлений градиентов алгоритм использует метрику косинуса разности квадратов, которую можно эффективно подсчитать, используя быстрые преобразования Фурье [2]. По результатам сравнения выбираются лучшие масштаб и сдвиг, и происходит вклейка.

Время работы алгоритма для текстур в несколько мегапикселей составляет от нескольких секунд до минуты при Matlab реализации. Ни один из

существующих алгоритмов не предназначен для восстановления сильно загроможденных текстур, поэтому не справляется с работой. По результатам экспериментов было показано, что предложенная нами схема позволяет качественно решать поставленную задачу в большинстве случаев.

### Литература

1. Liebowitz D., Criminisi A., Zisserman A. Creating Architectural Models from Images // EuroGraphics, No.3, 1999.
2. Paglieroni D., Eppler W., Resolution analysis for Gradient Direction Matching of object model edges to overhead images // Computer Vision and Image Understanding, v.2, 2009.
3. Park M., Collins R., Liu Y. Deformed Lattice Discovery Via Efficient Mean-Shift Belief Propagation // ECCV: Part II, 2009.

## Алгоритм автоматического совмещения аэрофотоснимков с малым перекрытием

*Кривовязь Г.Р., Мизотин М.М.*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,*

*Факультет вычислительной математики и кибернетики, Москва,*

*Россия*

*email: krigl@mail.ru*

Задача совмещения аэрофотоснимков представляет значительный научный и практический интерес, как сама по себе, так и в качестве промежуточного этапа процесса построения трехмерной модели местности. В связи с большим объемом исходных данных, становится актуальной разработка автоматического алгоритма, позволяющего исключить ручную работу оператора.

Являясь частным случаем задачи сопоставления изображений, совмещение аэрофотоснимков обладает рядом особенностей, одна из которых - потенциально малое (до 20%) перекрытие снимков. Наиболее универсальный и успешный подход к решению задач совмещения заключается в сопоставлении точечных особенностей (нами был использован детектор Харриса [2]), с помощью дескрипторов (например, SIFT [3]). Отбор правильных соответствий обычно производится устойчивыми алгоритмами типа RANSAC. Однако, вероятность нахождения ими правильного решения резко падает с уменьшением площади перекрытия снимков.

Предлагаемый алгоритм решает задачу в две стадии. Сначала с помощью преобразования типа Хафа, устойчивого к малому проценту правильных соответствий, определяется область перекрытия снимков. Затем производится сопоставление изображений только по данной области, что позволяет добиться надежного совмещения даже в сложных случаях.

Определение области перекрытия алгоритма также разбито на два этапа. На первом этапе производится определение относительного поворота снимков с помощью ориентаций характеристических точек, которые вычисляются при расчете дескриптора SIFT. На втором этапе определяется относительное смещение снимков с учетом уже известного поворота. Каждая пара точек голосует за свой сдвиг по  $x$  и по  $y$ . Аккумулятор сглаживается и в нем выделяется максимум, который приближается двумерной функцией Гаусса для уточнения координат пика и выделения правильных соответствий. Наконец, производится уточнение сдвига и поворота по выбранным точкам. Заметим, что отношение величины следующего пика к величине наивысшего может служить мерой надежности результата.

Было проведено сравнение предложенного алгоритма с альтернативной схемой поиска перекрытия, использующей алгоритм PROSAC [1] для оценки модели сдвига-поворота. По результатам синтетических тестов, а также экспериментов на реальных данных наш алгоритм превзошел альтернативный метод по числу правильно совмещенных аэрофотоснимков с малыми перекрытиями. К достоинствам предложенного метода также можно отнести скорость, детерминированность и отсутствие необходимости в настройке порога.

#### Литература

1. O. Chum, J. Matas. Matching with PROSAC – Progressive Sample Consensus // Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2005, Vol. 1, pp. 220-226.
2. C. Harris, M. Stephens. A combined corner and edge detector // Proceedings of The Fourth Alvey Vision Conference, Manchester, 1988, pp. 147–151.
3. D. G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints // International Journal of Computer Vision, Vol. 60, No. 2, 2004, pp. 91-110.

## Система стеганографии для видео последовательностей

*Куликов Дмитрий Леонидович*

*Кандидат наук*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,*

*Россия*

*email: [dyomich@yandex.ru](mailto:dyomich@yandex.ru)*

Задача стеганографии, т.е. сокрытия информации, известна давно, но для разных типов сигналов-носителей методы и алгоритмы различны [1], но существуют два основных подхода: изменение наименее значимого байта (LSB) и метод добавления.

1. Так для сокрытия дополнительной информации в тексте обычно используют дополнительные пробелы и символы табуляции в конце строк – это пример метода добавления.

2. Методы стеганографии в аудио основаны либо на LSB (но это обычно приводит к заметным для человека искажениям), либо на добавлении шума к аудио.

3. Для сокрытия информации в изображениях обычно применяют методы LSB – изменение яркости либо искажения коэффициентов ДКП (дискретное косинусное преобразование).

4. На современном этапе для стеганографии в видео уже более 10 лет в основном используют искажения ДКП коэффициентов[2].

Основная проблема существующих подходов к сокрытию информации в видео – это потеря или сильное искажение этой информации при типичных преобразованиях сигнала. Особенно это актуально, если известно, что в данном видео сокрыта информация. Так для одной из задач стеганографии в видео – «водяные знаки», когда, например, в фильм, транслируемый в кинотеатре, записывается (секретным невидимым для человека образом) номер пленки или номер кинотеатра, для того, чтобы по пиратской копии можно было понять, в каком кинотеатре была произведена запись. Но в таком случае при применении метода на основе ДКП-коэффициентов при естественном перекодировании видео и изменении яркости во время съёмки большая часть сокрытой информации будет искажена или потеряна.

Автором разработан метод для добавления информации в видео на основе подбора блоков и матричного изменения яркости и применении помехоустойчивого кодирования[3]. Данный метод позволяет решать задачу стеганографии в видео, при этом при типичных преобразованиях сигнала (кодирование, изменение яркости, контраста, размытие) скрытый текст может быть восстановлен достаточно успешно. На основе предложенного метода реализован прототип программной системы, позволяющий добавлять и извлекать из несжатого видео текстовую информацию с парольной защитой.

## Литература

1. Jonathan K. Su, Frank Hartung, Bernd Girod Digital Watermarking of Text, Image, and Video Documents // Computers & Graphics, 1998, volume 22, pp. 687-695
2. Chiou-Ting Hsu, Ja-Ling Wu. DCT-based watermarking for video// IEEE Transactions on Consumer Electronics, 1998, volume 44, pp. 206-216
3. Kulikov D., Putilin S. Low-Complexity Video Watermarking Using Mean Block Luminance Quantization. GraphiCon-2008, Moscow, 2008, pp. 302-303
4. L. Parameswaran, K. Anbumani. Content-Based Watermarking for Image Authentication Using Independent Component Analysis//Informatica 32, 299-306, 2008.

5. Mona F. M. Mursi, Ghazy M.R. Assassa, Hatim A. Aboalsamh, Khaled Alghathbar. A DCT-Based Secure JPEG Image Authentication. World Academy of Science, Engineering and Technology 53 2009, pp. 682-687

## **Физически аккуратное моделирование отражательной способности материалов поверхностей плоских объектов по фотоизображениям**

*Лебедев А.С., Ильин А.А.*

*Студент*

*МГУ им.Ломоносова, Вычислительной математики и кибернетики,  
Москва, Россия*

*email: andrewlebedev@yandex.ru*

При проектировании автомобилей, самолетов, зданий, помещений и т.п. обычно производится предварительный расчет освещения и визуальная оценка дизайнерских решений для контроля эстетичности внешнего вида. Для решения этих задач необходимо уметь моделировать реальные материалы (вручную или автоматически). При автоматическом восстановлении моделей материалов по фотоизображениям необходимо учитывать технические характеристики фотоаппарата, с помощью которого были получены снимки, значения выдержки, характеристики источника света. Также большое внимание уделяется цветовому пространству, в котором производится вычисление двулучевой функции отражения (ДФО) материала. Вместо пространства sRGBV используется пространство XYZ, что обеспечивает более точное восстановление материала (см. [1]). В настоящей работе описывается процесс восстановления физически аккуратной модели материала, предложенной нами ранее в [2], с учетом этих характеристик.

В данной работе предлагается следующий процесс восстановления модели материала с плоской поверхностью. В темном помещении с точечным источником света (с известными спектральными характеристиками) производится фотографирование плоской поверхности материала с разной выдержкой и с разных позиций. После этого производится автоматическое определение положения источников света и камеры по фотографиям. На следующем этапе производится расчет кривой отклика камеры (зависимость интенсивности точки изображения от реальной освещенности). В ряде работ предполагается, что кривая отклика имеет вид  $f = x^\gamma$ , что эквивалентно простейшей гамма-коррекции, либо является линейной функцией. Это приводит к неточностям при моделировании материала, т.к. реальные камеры имеют более сложную кривую отклика. Знание кривой отклика камеры позволяет получить значения освещенности по набору фотографий с разной выдержкой. По этим значениям и данным о спектральной кривой источника восстанавливается модель освещения. При визуализации для каждой точки изображения производится обратное преобразование рассчитанной освещенности в интенсивность с использованием той же кривой отклика.

## Литература

1. Michael Goesele, Hendrik P. A. Lensch, Hans-Peter Seidel. Validation of Color Managed 3D Appearance Acquisition // Proceedings of the 12th Color Imaging Conference Color Science and Engineering: Systems, Technologies, Applications. 2004. P. 265-270.
2. Ильин А.А., Лебедев А.С., Синяевский В.А., Игнатенко А.В. Моделирование отражательных свойств материалов плоских объектов по фотоизображениям // Proc. of Graphicon'2009. 2009. С. 198-202.

### Алгоритм устранения эффекта Гиббса при восстановлении сжатого мультипликационного видео

*Моисейцев Алексей Борисович*

*студент (специалист)*

*Московский государственный университет имени М.В. Ломоносова  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: himeraster@gmail.com*

Одной из основных проблем, возникающих при работе с цифровым видео, является проблема компактного хранения видеоданных. За последние два десятилетия были разработаны различные алгоритмы сжатия видео. Во многих случаях не требуется идентичность исходного и закодированного видео, поэтому большое распространение получили алгоритмы сжатия с потерями, так как они позволяют достигнуть существенно более высоких степеней сжатия по сравнению с беспотерными алгоритмами. Многие из современных методов сжатия используют в своей работе двумерные дискретные преобразования Фурье. При этом в сжатом видео коэффициенты, отвечающие за высокочастотные составляющие сигнала, сохраняются с пониженной точностью, либо не сохраняются вовсе. Это приводит к тому, что при высоких степенях сжатия в видео возникают артефакты, например эффект Гиббса (рингинг), блочность и др. Так эффект Гиббса особенно сильно заметен в областях изображения через которые проходят резкие границы. В силу того что современные алгоритмы сжатия, использующие квантование частотных коэффициентов, ориентированы на сжатие фотореалистичного видео, то при использовании таких алгоритмов для сжатия мультипликационного видео одним из наиболее заметных артефактов сжатия становится именно эффект Гиббса, а в связи с тем, что существует большое количество мультипликационного видео, сжатого такими кодеками, возникает задача устранения артефактов, вызванных им.

Разработано немало методов постобработки видео, однако большинство из них либо используют простые математические модели ориентированные на фотореалистичное видео и не учитывающие специфику видео мультипликационного [1, 2], либо используют сложные модели, трудноприменимые в приложениях реального времени [4].

Предлагаемый алгоритм позволяет обрабатывать видео в реальном времени при этом обеспечивая высокое качество фильтрации. Для фильтрации кадров видео в пространственной области был предложен многопроходный билатеральный фильтр — модификация билатерального фильтра [3], применяемого до этого только в алгоритмах фильтрации немультимпликативного видео. Многократное применение билатерального фильтра к изображению позволяет сохранять резкие границы на видео и вместе с тем удалять артефакты, вызванные эффектом Гиббса, которые возникают рядом с такими границами. Так же предлагаемый алгоритм для ускорения своей работы использует временную избыточность видео, исключая из фильтрации неменяющиеся во времени области кадра, наличие которых характерно для многих мультипликационных видеопоследовательностей.

Разработанный алгоритм показал высокое объективное (на основе метрики Y-PSNR) и визуальное качество в сравнении с другими методами [1,2,3,4]. Алгоритм реализован в виде библиотеки для программы обработки видео VirtualDub и доступен по адресу: [http://compression.ru/video/cartoon\\_restore/index.htm](http://compression.ru/video/cartoon_restore/index.htm).

#### Литература

1. Каур А., "Reduction of ringing noise in transform image coding using simple adaptive filter"// Electronics Letters, 1998, vol. 34, pp. 2110–2112.
2. H. S. Kong, Y. Nie, A. Vetro, H. Sun, K. E. Barner, "Coding artifact reduction using edge map guided adaptive and fuzzy filter"// IEEE International Conference on Multimedia and Expo, 2004, pp. II-1135–1138.
3. Tomasi C., Manduchi R., "Bilateral Filtering for Gray and Color Images"// Proceedings of the Sixth International Conference on Computer Vision, 1998, p. 839.
4. Guangyu Wang, Tien-Tsin Wong, Pheng-Ann Heng, "Deringing Cartoons by Image Analogies"// ACM Transactions on Graphics, 2006, vol. 25, no. 4, pp. 1360–1379.

### Система построения трехмерных моделей зданий по поэтажным планам

*Петрушкина Анастасия Александровна*

*Студент*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: dicentra13@gmail.com*

Задача построения трехмерной модели здания возникает в рамках создания информационной модели строения, которая используется для информационной поддержки процессов технического управления в течение всего жизненного цикла здания. При построении трехмерной модели здания зачастую

единственным источником информации о внутренней структуре помещений являются поэтажные планы. Процесс построения трехмерной модели здания по поэтажным планам, заданным в векторном формате, включает в себя следующие этапы:

1. Восстановление топологической структуры с необходимой геометрической коррекцией
2. Распознавание объектов на планах этажей (двери, окна)
3. Построение трехмерных моделей этажей
4. Совмещение моделей этажей в модель здания

Ключевым этапом является распознавание элементов чертежей: выделение групп линий чертежа, образующих объекты определенного типа. Данная задача сводится к задаче распознавания образов со структурными ограничениями, поэтому для ее решения целесообразно использовать структурные методы распознавания образов.

Одним из перспективных подходов к решению задачи распознавания элементов чертежей является сетевой метод, основанный на поиске изоморфных подграфов с применением подхода „разделяй и властвуй“, предложенный в работе [1]. Структура чертежа и шаблонов элементов описывается в виде графов отношений атрибутов (attributed relational graph), и в графе данных выполняется поиск всех подграфов, изоморфных графам шаблонов. Для учета погрешностей входных данных вводится мера сходства графов. Авторами была показана применимость данного подхода к решению проблемы определения элементов инженерных чертежей [2], при этом на структуру чертежа накладывается ряд ограничений, в общем случае не выполняющихся для поэтажных планов.

В данной работе показывается применимость сетевого метода поиска изоморфных подграфов к решению задачи распознавания элементов поэтажных планов и рассматривается архитектура экспериментальной системы построения трехмерной модели здания. Система поддерживает полный цикл построения трехмерной модели: от восстановления структуры исходных чертежей до визуализации полученного результата. Также обеспечивается возможность пользовательского задания шаблонов объектов, распознавание которых производится на основе рассматриваемого сетевого подхода. Сетевой алгоритм поиска изоморфных подграфов реализован в виде библиотеки, которая может быть использована для применения данного алгоритма и в других предметных областях.

### Литература

1. В. Т. Messmer, Н. Bunke. A network based approach to exact and inexact graph matching. Technical Report IAM-93-021, Universits Bern, September 1993.



2. B. T. Messmer, H. Bunke. Automatic Learning and Recognition of Graphical Symbols in Engineering Drawings. // In R. Kasturi and K. Tombre (Eds.), Graphics Recognition—Methods and Applications, Volume 1072 of Lecture Notes in Computer Science, pp. 123–134. Springer-Verlag, May 1996.

## Восстановление параметрических моделей отражательных свойств материалов по двум фотоизображениям

*Синяевский В.А., Ильин А.А.*

*Студент*

*МГУ им.Ломоносова, Вычислительной математики и кибернетики,  
Москва, Россия*

*email: zorro5101@mail.ru*

В настоящее время одной из актуальных задач в области компьютерной графики является автоматизация процесса восстановления моделей отражательных свойств реальных материалов ([1], [2]). Некоторые ее решения связаны с использованием высокоточных, но дорогих установок, восстанавливающих отражательные свойства материалов с высокой степенью точности. Другие основаны на определении моделей отражательных свойств по фотоизображениям. Преимуществами этих решений являются быстрое получение искомым моделей и возможность его применения в бытовых условиях.

Обычно модели отражательных свойств материалов задаются сеточными или параметрическими функциями. Сеточное задание позволяет охватить широкий класс материалов, однако затратно по памяти. Параметрическое задание имеет более узкий класс моделируемых материалов, но такое задание не затратно по памяти и может быть быстро визуализировано на графическом процессоре (GPU).

В данной работе предложен алгоритм восстановления параметрических моделей материалов по паре фотоизображений. Алгоритм состоит из трех этапов – получение необходимых входных данных, предобработка изображений и восстановление модели отражательных свойств материала. На вход алгоритму задаются два фотоизображения объекта, поверхность которого покрыта исследуемым материалом. Одно из изображений сделано без вспышки, другое со вспышкой. Кроме того, на вход подаются EXIF-параметры фотоизображений (экспозиционные и фокальные параметры), и задается расстояние от камеры до исследуемого объекта. Далее происходит предобработка входных данных. Сглаживание изображений позволяет выделить текстурную карту и отделить ее от материала. Для совместной работы с пикселями на двух фотоизображениях требуется изображения привести их в один яркостной диапазон. Для этого применяется коррекция изображений по кривой отклика фотокамеры, а также их нормализация по экспозиции. Используя полученные изображения, приведенные в один яркостной диапазон, можно отделить с изображения со вспышкой ambient-составляющую материала.

На заключительном этапе происходит определение коэффициентов параметрической модели. Используя обработанные входные изображения, набор параметров модели отражательных свойств материала определяется методом Левенберга-Маркардта.

### Литература

1. Волобой А.Г., Галактионов В.А., Ершов С.В., Летунов А.А., Потемин И.С. Аппаратно-программный комплекс для измерения светорассеивающих свойств поверхностей // Информационные технологии и вычислительные системы, № 4. 2006.
2. P.Sikachev, A.Ilyin, A.Ignatenko. User-Assisted Acquisition, Processing and Rendering of Materials from Images // Proc. of Graphicon'2007. 2007. Moscow. pp. 131-134.

## Синтез HDR изображений при помощи технологии CUDA

*Скрябин Виктор Сергеевич*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,*

*Факультет вычислительной математики и кибернетики, Москва,*

*Россия*

*email: skrvikser@rambler.ru*

Широкий динамический диапазон яркостей, когда соотношение самой яркой области в поле зрения с самой темной составляет сотни тысяч или даже миллионы раз, без труда воспринимается человеческим глазом [1]. Однако возможностей современных графических дисплеев недостаточно для того, чтобы отобразить всю широту динамического диапазона реальной сцены, поскольку соотношение самой яркой области с самой темной не превышает у них сотен раз.

В последнее время развитие компьютерной графики все больше внимания уделяет изображениям широкого динамического диапазона (HDRI, High Dynamic Range Image). HDRI содержат значения яркости для каждой точки, выраженные вещественным числом, а значит, их диапазон также может составлять миллионы раз. В настоящий момент существуют различные способы получения HDRI реальной сцены, но одной из наиболее распространенных является обработка фотографий с различной экспозицией [3].

Рассматриваемый процесс требует от компьютера значительных временных ресурсов, поскольку необходимо учитывать значение каждой точки на каждой фотографии из набора. Однако существуют задачи, в которых синтезировать HDRI необходимо в реальном времени или с достаточной степенью интерактивности. Например, при улучшении качества изображения с цифрового микроскопа [2]. Зачастую при использовании центрального процессора (CPU) для обработки данных невозможно добиться того, чтобы изображение

формировалось в реальном времени. Особенности архитектуры современных графических процессоров (GPU) позволяют ускорить процесс синтеза HDRi на основе совместного использования CPU и GPU в гетерогенной вычислительной модели [4]. Суть данной модели заключена в том, что последовательные блоки алгоритма выполняются на CPU, а вычислительно-тяжелая часть на GPU. Вычисления на GPU могут быть выполнены с помощью технологии параллельных вычислений CUDA графических процессоров фирмы NVIDIA. Архитектура CUDA состоит из сотен процессорных ядер, работающих в связке и позволяющих разом справиться с набором данных, что естественно ускоряет процесс получения HDRi.

### Литература

1. Волобой А.Г., Галактионов В.А., Копылов Э.А., Шапиро Л.З. Моделирование естественного дневного освещения, задаваемого изображением с большим динамическим диапазоном // Программирование. 2006. No. 5. С. 62–80.
2. Матросов М., Гаганов В., Игнатенко А., Сивоволенко С. Улучшение качества изображения с микроскопа при помощи технологии HDRi в интерактивном режиме // Труды конференции ГрафиКон'2009. М., 2009. С. 332–335.
3. Debevec P., Malik J. Recovering High Dynamic Range Radiance Maps from Photographs // Computer Graphics (Proc. ACM SIGGRAPH'97). 1997, p. 369–378.
4. Что такое вычисления на GPU?: [www.nvidia.ru/page/gpu\\_computing.html](http://www.nvidia.ru/page/gpu_computing.html)

## Графическая модель для оценки геометрии городских сцен по одному изображению

*Третьяк Елена Викторовна*

*Студент*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,*

*Россия*

*email: tretiak.elena@gmail.com*

Задачей анализа геометрии городских сцен по фотографиям занимаются уже давно. Наиболее известный подход основан на использовании геометрических подсказок, таких как группы параллельных прямых, образованных стенами, окнами и т.д. Такие группы параллельны в пространстве, но на фотографии пересекаются из-за свойств перспективной проекции. Точки пересечения этих групп называются точками схода.

На фотографиях городских сцен в основном присутствуют группы вертикальных и горизонтальных отрезков. Точка схода группы вертикальных

прямых называется зенитом, остальные точки схода лежат на одной прямой – горизонте. Задача определения горизонта и зенита применяется для стабилизации камеры, выравнивания фотографий, определения относительного размера объектов и т.д. Существующие методы обычно находят три точки схода, относящиеся к взаимно перпендикулярным направлениям, либо все точки без учёта их взаимного расположения. Наш метод использует промежуточный подход: определяется любое количество точек схода, но при этом учитывается, что горизонтальные точки схода лежат на одной прямой, а одна точка относится к зениту

В основе нашего метода лежит графическая модель, использующая в качестве входных данных вычисленные на этапе инициализации отрезки и кандидаты в точки схода. Каждому кандидату в точку схода соответствует случайная величина, принимающая значения 1, если точка схода включена, иначе – 0; отрезкам – случайные величины, принимающие значения 0, 1, 2..n, где n – количество кандидатов в точки схода; 0 обозначает, что отрезок относится к фоновому шуму, значение  $i$ , отличное от 0, обозначает принадлежность к точке схода с номером  $i$ , что возможно только в случае, если  $i$ я точка включена, тогда к энергии системы добавляется штраф принадлежности отрезка к  $i$ й точке. Также вводится штраф для каждой пары включённых горизонтальных точек схода, пропорциональный углу наклона прямой, проходящей через эти точки. Ответ находится с помощью минимизации энергии модели с использованием метода имитации отжига, который довольно быстро сходится и работает всего несколько секунд на сложных изображениях.

Для тестирования мы использовали 2 размеченные базы. Первая – York Urban[1] содержит простые изображения с 3мя точками схода, вторая была составлена нами и включает в себя более сложные изображения. Метод сравнивался с одним из последних подходов [2] с помощью метрики качества нахождения горизонта (расстояние в пределах картинки между вычисленным горизонтом и ground truth). Тестирование показало, что использование априорных предположений о расположении горизонта позволяет более точно определять точки схода и линию горизонта. Средняя ошибка нахождения горизонта по обеим базам для нашего метода оказалась в несколько раз меньше, чем для предложенного в статье[2].

### Литература

1. P. Denis, J. H. Elder, F. J. Estrada. Efficient Edge-Based methods for estimating manhattan frames in urban imagery//ECCV'08
2. J.-P. Tardif, Non-Iterative Approach for Fast and Accurate Vanishing Point Detection//ICCV'09

### Слова благодарности

Автор выражает благодарность Бариновой О.С. Работа выполнена при поддержке Microsoft Research

# Конвейер калибровки и текстурирования модели головы по одной фотографии

*Федюков Максим Александрович*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,*

*Россия*

*email: maxim.fedyukov@gmail.com*

Задача моделирования головы человека является актуальной во областях видеокодирования на базе трехмерных моделей, моделирования виртуального присутствия и других. В работе представлен конвейер калибровки параметризованной полигональной модели головы человека, а также ее текстурирования. Параметризованные полигональные модели являются более общим случаем моделей активной формы [1] и частью трехмерных морфируемых моделей [2]. Универсальная разметка  $T := (t_1, t_2, \dots, t_M)$ , определяемая в начале работы конвейера и используемая на последующих этапах, представляет собой набор векторов  $t_i := (t_1^i, t_2^i, \dots, t_{L_i}^i)$ , элементами которых являются индексы  $t_i^i$  вершин базовой полигональной модели. Первый этап — разметка черт лица на входной фотографии, осуществляемая с помощью модели активной формы. Следующий этап — подгонка фотографий. Фотографии располагаются в пространстве, корректируются их положение, поворот и масштаб. Следующий этап — построение параметрической модели. Задача сводится к поиску значений из наборов параметров  $p^m$ , минимизирующих функционалы  $G_m = \sum_{i=l_m}^{l_m+1} \sum_{j=1}^{L_i} (\omega_j^{k_i} - \omega_j^{k_i'})^2$ . Следующий этап — подгонка полигональной модели. Подгонка происходит итеративно, на каждой итерации  $i$  вершины, принадлежащие одной черте лица (заданной универсальной разметкой  $T$ ), перемещаются в соответствующие координаты  $\hat{Q}$ :  $\forall i, j, \tilde{c}_{i_j} = \hat{q}_j^i$ . Веса перебираемых вершин вычисляются следующим образом:  $\nu_{\tilde{c}_{i_j}}^k = \max(\frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{\lambda j}{2\sigma^2}), \nu_{\tilde{c}_{i_j}}^k)$ , где  $\sigma$  — коэффициент масштаба,  $\lambda$  — нормировочный коэффициент. Следующий этап — генерация текстуры. Для генерации текстуры создается текстурированная полигональная модель  $\tilde{d}$ , значения элементов которой задаются следующим образом:  $\tilde{d}_x = 0$ ,  $\tilde{d}_y = \tilde{c}_u$ ,  $\tilde{d}_z = \tilde{c}_v$ ,  $\tilde{d}_u = \mathbf{P}\tilde{c}_y$ ,  $\tilde{d}_v = \mathbf{P}\tilde{c}_z$ , где  $\mathbf{P}$  — оператор ортогонального проецирования. Полученная плоская полигональная модель визуализируется в текстуру. Следующий этап — фильтрация текстуры. Текстура дублируется, исходная обрабатывается билатеральным фильтром, дублированная — фильтром Гаусса. На исходную накладывается маска прозрачности, каждый пиксель которой определяется косинусом угла между нормалью и направлением на камеру, и текстуры смешиваются.

## Литература

1. C. Sun, M. Xie, An enhanced Active Shape Model for facial features extraction // IEEE ICCT, 2008. P. 661-664.
2. B.N. Kang, H. Byun, D. Kim, Multi-resolution 3D Morphable Models and

## **Адаптивная трассировка путей на графических процессорах**

**Фролов Владимир Александрович**

*Студент*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: vova@frolov.pp.ru*

### **Трассировка путей**

Трассировка путей - это метод вычисления интеграла освещенности, позволяющий моделировать практически любые физические явления, обусловленные поведением света. В самом общем случае трассировка путей представляет собой вычисления интеграла освещенности при помощи метода Монте-Карло [1].

### **Предлагаемый алгоритм**

Тривиальная реализация алгоритма не была бы эффективна, так как графические процессоры (GPU) имеют определенные особенности. Одна из проблем связана с тем, что объем работы в различных частях изображения разный. Предлагается адаптивный алгоритм, разделяющий изображение на блоки и обрабатывающий одновременно массивы блоков. В начале процесса синтеза изображения мы имеем достаточно блоков, чтобы загрузить GPU вычислениями в режиме один луч на пиксел. С течением времени остается лишь небольшое число блоков, требующих вычислений, но количество работы для GPU не уменьшается за счет увеличения числа лучей, приходящихся на один пиксел. В конце концов все вычислительные ресурсы GPU будут сосредоточены на наиболее сложных участках сцены.

### **Литература**

1. Pharr, M. and Humphreys, G. 2004 Physically Based Rendering: from Theory to Implementation. Morgan Kaufmann Publishers Inc

### **Иллюстрации**

## **Автоматический анализ и обработка снимков пищеварительного тракта, сделанных в процессе капсульного эндоскопического исследования**

**Хабидуллин Т.Р., Михайлов Д.М.**

*Студент*

*Московский инженерно-физический институт, Факультет  
кибернетики, Москва, Россия*

*email: timur.khab@gmail.com*

Визуальный осмотр желудочно-кишечного тракта является одной из наиболее распространенных медицинских процедур. Капсульная эндоскопия —

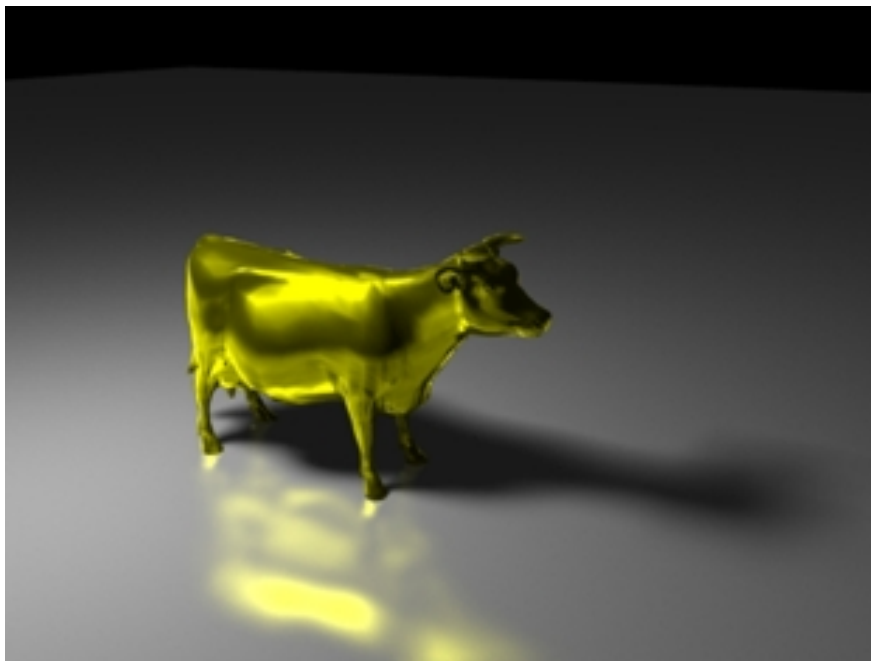


Рис. 1. Время синтеза изображения - 16 секунд на GTX260 в разрешении 1024x768.

новый, комфортный метод обследования, заменяющий две неприятные для пациента процедуры — гастроскопию (введение трубки с видеокамерой через рот) и колоноскопию (введение трубки через задний проход). В случае капсульной эндоскопии, пациент проглатывает миниатюрную капсулу, которая движется естественным образом по желудочно-кишечному тракту. Она производит постоянную фотосъемку и передает данные по радиоканалу на устройство хранения, которое пациент в течении процедуры носит на специальном ремне. Капсула выводится из организма естественным образом.

Классическая гастроскопия обладает рядом осложнений, в основном, — со стороны дыхательной и сердечно-сосудистой системы (30общего числа осложнений соответственно [1]), которые вызваны возникающим рвотным рефлексом. Капсула не провоцирует данных осложнений. Будучи однократно используемой, капсула также не может служить источником передачи инфекции от пациента к пациенту.

Капсульная эндоскопия не только заменяет гастро- и колоноскопию, но и является единственным методом не инвазивного осмотра тонкого кишечника.

ка [3] — самого длинного органа пищеварительного тракта [2]. Кроме того, длительное нахождение капсулы в желудочно-кишечном тракте значительно повысит вероятность обнаружения скрытых кровотечений [4].

Представленная работа — часть разработки общего комплекса капсульной эндоскопии, включающая создание специфических алгоритмов и программного обеспечения для автоматического анализа изображений.

Поскольку полное исследование занимает 7-10 часов, врач-эндоскопист тратит значительное время и усилия на анализ полученных снимков [5]. Существенно повысить эффективность работы врача (разумеется, не исключая ее полностью) позволит автоматическая обработка изображений, состоящая из

удаления повторяющихся изображений (поскольку на некоторых этапах пищеварения капсула может находиться неподвижно значительное время)

опознания и выделения заболеваний с учетом анамнеза (истории болезни) и предварительного заключения врача

подстройки параметров изображений (яркость, контраст, цветовой баланс) с обнаружениями, которые позволят четко их обозначить

Подобные алгоритмы пока не реализованы ни в одном из комплексов капсульной эндоскопии и созданное на их основе программное обеспечение является уникальным.

### Литература

1. Блок Б., Шахшаль Г., Шмидт Г. Гастроскопия, М. 2007
2. Питер Р. М. Секреты гастроэнтерологии, М. 2008
3. Саблин О., Гриневич В., Успенский Ю., Ратников В. Функциональная диагностика в гастроэнтерологии, Санкт-Петербург 2002
4. Faigel D., Cave D. Capsule Endoscopy, Elsevier Inc. 2008
5. Halpern M., Jacob H. Atlas of Capsule Endoscopy, Given Imaging 2002

### О подходе к объективной оценке качества цветных цифровых изображений

*Хорунжий Михаил Дмитриевич*

*Аспирант*

*Санкт-Петербургский Государственный Университет Кино и Телевидения, приборов систем кино и телевидения, Санкт-Петербург, Россия*

*email: horungij@gmail.com*

Объективная оценка качества цветных изображений в автоматическом режиме является важной прикладной задачей, решение которой позволит оптимизировать параметры электронного кинопоказа. Этап восстановления колориметрических свойств контента киноплёнок с помощью фильмосканера,



требует более гибкого регулирования различного семантического содержания при экранной проекции. В качестве модели управления качеством цветных изображений предлагается рассмотреть систему линейных интегральных уравнений второго рода с тремя неизвестными функциями для канала яркости и двух каналов цветности:

$$\begin{cases} y_{unknown1}(x) - c_1 \int_0^L H_{e1}(w) \min_Q f_1(L - L(Q)) dQ = \delta K_1(\varphi) \\ y_{unknown2}(x) - c_2 \int_{-a}^a H_{e2}(w) \min_Q f_2(a - a(Q)) dQ = \delta K_2(\varphi) \\ y_{unknown3}(x) - c_3 \int_{-b}^b H_{e3}(w) \min_Q f_3(b - b(Q)) dQ = \delta K_3(\varphi) \end{cases}$$

, где  $y_{unknown1}(x)$ ,  $y_{unknown2}(x)$ ,  $y_{unknown3}(x)$  - искомые функции, обеспечивающие оптимальные значения цветовых характеристик;  $c_1, c_2, c_3$  - весовые коэффициенты; энтальпия энтропия характеризует полосовой фильтр на входе  $H_e(w) = \frac{1}{2\pi i} \int_0^{pi} \frac{\ln p(x)}{2\pi i} dx$ , где  $p(x)$  - плотность распределения вероятности появления сообщения;  $\min_Q f_1(L - L(Q)), \min_Q f_2(a - a(Q)), \min_Q f_3(b - b(Q))$  - минимальное различие яркости и цветовых составляющих эталонного и искаженного изображений;  $\delta K_1(\varphi), \delta K_2(\varphi), \delta K_3(\varphi)$  - абсолютная разность функций передачи модуляции эталонного и искаженного изображений для каждого канала, где  $\varphi$  - пространственная частота. Решение приведенной системы уравнений, можно представить в общем виде:  $y_{\Sigma}(x) = \delta K_{\Sigma}(\varphi) + c_{\Sigma} \int_{-1}^1 R(\varphi, w) f_{\Sigma}(\delta L, \delta a, \delta b) d\varphi$ . Нами эмпирически были установлены оптимальные значения вносимых искажений колориметрических характеристик и шумов в ряде исследуемых оцифрованных изображений. Моделирование процесса восстановления эталонных значений изображений производилось в среде MATLAB. Согласование результирующих коэффициентов хроматических искажений по цветовому тону и насыщенности ( $\eta_{\Sigma}, \eta'_{\Sigma}$ ), полученных при приборных оценках сегментированного цветного экранного изображения, с мерикой цветовых различий  $\delta E_{ab}$  показало степень корреляции равную 0,8 [1]. Таким образом предложенная модель может быть внедрена в соответствующий микроконтроллер для операции цветокоррекции в фильсканере. Методика для исследования качества восстановленных цветных изображений успешно используется в учебном процессе СПбГУКиТ.

### Литература

1. Хорунжий М.Д. Анализ искажений цветопередачи экранных изображений посредством пространственного приложения модели LAB// Мир Техники Кино. М, 2009. №13. С. 4-11.

# Ускорение обработки видеоданных с применением современных графических процессоров

*Цветков Илья Игоревич*

*Студент (специалист)*

*Московский государственный университет имени М. В. Ломоносова,  
факультет Вычислительной математики и кибернетики, Москва,  
Россия*

*email: tsvetkov.ilya@gmail.com*

Использование цифрового представления видеоинформации в значительной мере расширяет возможности по работе с видео, позволяя использовать для этого весь спектр алгоритмов обработки данных при помощи вычислительной техники. Повышение разрешения видео и усложнение алгоритмов его обработки неизбежно влечёт за собой рост требуемых вычислительных ресурсов. Отличительной особенностью большинства алгоритмов обработки видеопоследовательностей является возможность их эффективной реализации на процессорах с параллельной архитектурой.

На данный момент наиболее широко распространёнными процессорами с параллельной архитектурой являются современные программируемые графические процессоры. Их высокая производительность и широкая распространённость стали основой для проведения множества научных работ в области реализации алгоритмов на графических процессорах [2]. Значительная часть этих работ посвящена алгоритмам обработки изображений и видео, например, оценке движения в видео [1], шумоподавлению с применением билатеральной фильтрации [3].

В данной работе решается задача повышения производительность ресурсоёмких алгоритмов обработки видеопоследовательностей за счёт использования вычислительных ресурсов графического процессора. В отличие от большинства других работ в этой области, для реализации были выбраны наиболее качественные алгоритмы обработки видео, обладающие высокими требованиями к вычислительным ресурсам.

В рамках данной работы было проведено исследование методов применения современных графических процессоров для обработки видео. В результате, в качестве интерфейса взаимодействия с графическим процессором была выбрана библиотека OpenGL [4]. На графическом процессоре было реализовано несколько алгоритмов подавления шума в видеопоследовательностях, а также высококачественный алгоритм повышения чёткости видео. Все реализованные алгоритмы были протестированы в сравнении с реализациями, выполняющими вычисления на центральном процессоре. Тестирование показало, что использование графического процессора позволяет ускорить обработку видео от 5 до 15 раз без потерь в качестве.

## Литература

1. C.-W. Ho, O. Au, S.-H. Chan, S.-K. Yip, H.-M. Wong. Motion Estimation for H.264/AVC using Programmable Graphics Hardware // IEEE International

Conference on Multimedia Expo (ICME), pp. 2049-2052, 2006.

2. John D. Owens, David Luebke, Naga Govindaraju, Mark Harris, Jens Kruger, Aaron E. Lefohn, Timothy J. Purcell. A Survey of General-Purpose Computation on Graphics Hardware // Computer Graphics Forum, Volume 26, Number 1, pp. 80-113, 2007.
3. Ivan Viola, Armin Kanitsar, Meister Eduard Groller. Hardware-Based Nonlinear Filtering and Segmentation using High-Level Shading Languages // IEEE Visualization, 2003.
4. OpenGL: <http://www.opengl.org/>

## Калибровка камеры по нескольким шаблонам

*Четвериков Никита Сергеевич*

*Студент*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: nikitachetv@mail.ru*

Калибровка камеры представляет собой определение её внутренних (фокусные расстояния, координаты принципиальной точки) и внешних (положение и направление камеры относительно сцены) параметров по фотографиям. Автоматическая калибровка камеры используется в задачах 3D-реконструкции по изображениям и фотограмметрии.

Данная задача обычно решается либо с помощью автокалибровки, например, путём сопоставления точечных особенностей на изображениях, либо с помощью калибровки по шаблону [1]. Калибровка по шаблону даёт лучший по точности и надёжности

результат, но обладает недостатками: ограниченный диапазон ракурсов для съёмки, и недостаточную для некоторых задач точность.

Для решения этих проблем предлагаемый алгоритм выполняет калибровку по произвольному числу шаблонов, при этом на положение шаблонов в пространстве ограничений не накладывается. Единственное условие - это связность графа

$G=(V,E)$ , где каждому шаблону соответствует вершина в  $V$ , а каждой фотографии с шаблонами  $S = (a_{k_1}, a_{k_2}, \dots, a_{k_n})$  - дуги  $(a_{k_i}, a_{k_j}) \in E, i \neq j$  между всеми возможными парами элементов из  $S$ . Шаблон представляет собой плоскую поверхность, например, лист бумаги, с нанесённым на неё рисунком в виде шахматной доски. Углы клеток шахматной доски распознаются автоматически.

Задача калибровки формулируется в виде единой задачи минимизации суммарной ошибки репроекции точек всех шаблонов на все изображения, на которых соответствующие шаблоны были найдены. Для её решения используется алгоритм

Левенберга–Марквардта [2]. Репроекция получается путём расчёта 3D-координат точек шаблона и последующей их проекции обратно на фотографию.

С помощью специального алгоритма выбирается базовый шаблон, с ним связывается базовая система координат: плоскость  $z=0$  совпадает с плоскостью шаблона, ось  $z$  направлена вверх перпендикулярно плоскости шаблона. Далее для каждого оставшегося шаблона  $i$  вводится матрица  $H_i$  перехода к его плоскости от базовой системы координат.

Таким образом, к стандартному набору параметров, используемому в существующих алгоритмах, добавляются матрицы перехода между шаблонами.

### Литература

1. Z. Zhang, A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11):1330-1334, 2000.
2. Ф. Гилл, Ю. Мюррей, М. Райт. Практическая оптимизация. М.: «Мир», 1985

## Классификация трёхмерных облаков точек с помощью неассоциативных Марковских сетей

*Шаповалов Роман Викторович*

*Студент*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: shapovalov@graphics.cs.msu.su*

В последнее время задача классификации лазерных сканов стала очень актуальна. Например, анализ данных, получаемых с лазерных сенсоров, незаменим для навигации мобильных роботов. Лазерная съёмка применяется также для создания моделей зданий в архитектуре, для обработки данных аэрофотосъёмки в геодезии, для контроля качества продукции, а также для сохранения культурного и исторического наследия. В отличие от фотографий, облако точек представлено в реальном масштабе и точнее отражает трёхмерную структуру сцены.

Для решения поставленной задачи мы разработали алгоритм машинного обучения, состоящий из двух стадий: обучения классификатора и классификации. Для обучения необходима выборка из размеченных вручную облаков точек. Обученный классификатор может быть использован для классификации тестовых сканов. На каждой из стадий строится граф над точками, рассчитываются локальные признаки точек и признаки для рёбер графа. На последнем этапе стадии обучения обучаются классификаторы, стадии классификации – производится вывод в Марковской сети.

Большинство существующих методов классификации облаков точек используют вывод в ассоциативной Марковской сети [1]. Минимизируется функция энергии Поттса. Ассоциативная Марковская сеть позволяет сглаживать результаты классификации. Существенным недостатком такой модели является неспособность задавать пространственные отношения между классами, такие как "дерево находится выше земли". Мы используем произвольный вид парных потенциалов, осуществляя приближённый вывод в Марковской сети с помощью алгоритма TRW-S [2]. Для настройки парных потенциалов используется наивный байесовский классификатор. Результат работы метода приведён в таблице.

Наш метод	земля	здание	автомобиль	дерево	куст	Recall	AMN recall
земля	398191	625	5680	7194	28299	0.9050	0.9263
здание	1199	14949	904	358	2182	0.7630	0.2078
автомобиль	1738	77	832	6	582	0.2572	0.0000
дерево	17143	7441	104	440946	66218	0.8291	0.9716
кусты	1391	229	152	97	5889	0.7591	0.4729
Precision	0.9488	0.6410	0.1084	0.9829	0.0571		
AMN Precision	0.9175	0.9995	0.0000	0.9501	0.3569		

Таблица 1. Результаты метода на лазерном скане "GML airborne dataset A", снятом с самолёта. Матрица ошибок для нашего метода, а также сравнение его с AMN с постоянными потенциалами

## Литература

1. D. Anguelov, B. Taskar, V. Chatalbashev, D. Koller, D. Gupta, G. Heitz, and A. Ng, "Discriminative learning of markov random fields for segmentation of 3-d scan data," In IEEE Conference on Computer Vision and Pattern Recognition, 2005.
2. V. Kolmogorov, "Convergent tree-reweighted message passing for energy minimization," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, 2006, p. 1568.

# Машинное обучение и анализ данных

## Автоматическое выявление терминологических вариантов в русскоязычных текстах

Антонов В.Ю., Ефремова Н.Э.

Студент

Московский государственный университет имени М.В. Ломоносова,

ВМиК, Москва, Россия

email: avadin@gmail.com

Выявление **терминов** — слов и словосочетаний, называющих понятия предметной области, — является важной задачей автоматической обработки научно-технических текстов. Одна из сложностей автоматического выявления терминов связана с тем, что при употреблении в текстах термины изменяются — усекаются, сокращаются и т.д.: *дисковый контроллер* — *контроллер диска*. Такие **терминологические варианты** представляют собой различные формы выражения одного понятия, поэтому их учет позволит получить более полную информацию о терминах и их употреблении.

Данная работа посвящена решению задачи выявления терминологических вариантов для русскоязычных текстов. Предлагаемый подход базируется на классификации, учитывающей типичные варианты употребления терминов в текстах на русском языке [1].

В большей части терминологических вариантов наблюдается изменение лексического состава термина и синтаксических связей внутри него. Для выявления таких вариантов были предложены методы, опирающиеся на правила варьирования терминов и их морфосинтаксическую структуру. При выявлении учитывается указанная информация и эвристики, помогающие определять, являются два словосочетания терминологическими вариантами или нет.

Для задания правил варьирования русскоязычных терминов нами был использован формальный язык LSPL, позволяющий описывать слова и словосочетания в виде лексико-синтаксических шаблонов. Например, шаблон

$$A1 N1 \langle A1=N1 \rangle \# A1 N2 \langle \text{Syn}(N1, N2), A1=N2 \rangle$$

задает правило варьирования термина с морфосинтаксической структурой прилагательное + существительное. “Левая часть” шаблона (до #) описыва-

ет структуру термина, “правая часть” задает его терминологический вариант. Конструкция *Sup(N1, N2)* означает, что существительные N1 и N2 должны быть синонимами. При использовании словаря, в котором слова *лучи* и *излучение* зафиксированы как синонимы, словосочетания *рентгеновские лучи* — *рентгеновское излучение* будут распознаны с помощью такого шаблона как терминологические варианты.

Нами разработаны программные средства, поддерживающие формирование по “правой части” правил варьирования возможных терминологических вариантов. При условии распознавания в тексте термина с помощью специальной процедуры формируются новые шаблоны, описывающие его терминологические варианты. Далее эти шаблоны используются для поиска среди словосочетаний текста.

Разработанные на языке C++ средства опираются на программную реализацию языка LSPL. Они апробированы на наборе научно-технических текстов, анализ полученных результатов показал состоятельность подхода.

### Литература

1. Большакова Е.И., Васильева Н.Э. Терминологическая вариантность и ее учет при автоматической обработке текстов // 11 Национальная конференция по искусственному интеллекту с международным участием КИИ-2008. Труды конференции в 3-х томах. Т. 2. – М.: Физматлит, 2008, с.174-182.

## Применение нейронных сетей для построения механических торговых систем.

*Глекин Александр Олегович*

*Студент*

*Московский физико-технический институт, Факультет инноваций и  
высоких технологий, Москва, Россия*

*email: graed@mail.ru*

Механические торговые системы или просто торговые роботы являются неотъемлемой частью сегодняшней биржевой торговли, что делает вопрос о технологии создания таких роботов чрезвычайно актуальным.

Целью данной работы является изучение возможности построения торговых роботов с использованием нейронных сетей. Торговый робот рассматривается как классификатор, на вход которого подаются исторические данные за определенный интервал времени, а на выходе возвращается класс, к которому принадлежат эти данные. Рассматривается три возможных класса: “buy”, “sell” и “nothing”.

С этой целью была создана компьютерная программа, реализующая следующую функциональность:

- графическая система для отображения рыночных данных;
- инструменты для изучения поведения роботов;
- инструменты для задания обучающих примеров и обучения сетей.

В работе были рассмотрены следующие нейросетевые архитектуры[1,2,3]:

- сети Хопфилда;
- сети Кохонена;
- многослойные сети;
- сети адаптивного резонанса.

Для тестирования роботов на основе нейронных сетей был создан симулятор рыночной ситуации, которая имеет место в реальном рынке. Эксперименты показали, что нейронные сети вполне успешно справляются с задачей распознавания рыночных ситуаций, что позволяет на их основе создавать торговых роботов для нахождения точек входа в рынок.

Так же в работе рассматривается вопрос о нахождении оптимальной стратегии выхода из рынка по заданным точкам входа.

### Литература

1. Головкин В. А. Нейронные сети: обучение, организация и применение. ИПРЖР, 2002.
2. Anthony M. Bartlett P.L. Neural Network Learning: Theoretical Foundations. Cambridge University Press 1999.
3. Heaton Research: <http://www.heatonresearch.com>

### Классификация прецедентов, образующих связанные множества, на примере данных лазерной локации

*Гусейнов Алексей Тимурович*

*Студент*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: kiberigus@gmail.com*

В докладе рассматриваются особенности применения методов классификации к задачам, в которых прецеденты с одинаковым классом образуют связанные множества. Как правило, отношение связности между прецедентами возникает в силу того, что они являются результатами измерений непрерывных величин в узлах некоторой сетки. Как следствие, класс является свойством не отдельного прецедента, а некоторой связанной области исходного пространства. Примером такой задачи является классификация данных лазерной локации. В этом случае исходное пространство состоит из ландшафта, яркости и цвета поверхности Земли.

Для решения таких задач чаще всего применяются методы, полагающие классы прецедентов независимыми величинами[1,3], что не соответствует действительности. С целью обеспечения корреляции классов соседних прецедентов признаковое пространство формируется их значений, вычисляемых по множеству близлежащих прецедентов. Альтернативным подходом является



использование методов классификации зависимых прецедентов, например сетей Маркова[2]. Но в силу неизвестности степени корреляции классов соседних прецедентов, используются довольно грубые приближения, приводящие к подавлению мелких деталей.

Для преодоления этих проблем предлагается новый подход к решению данной задачи, состоящий в попеременной классификации и уточнении структуры зависимостей. Первичная классификация дает представление о расположении и форме областей в исходном пространстве. Эти данные используются для оценки ковариации классов прецедентов либо введения специальных признаков. Кроме того, становится возможным учет априорных знаний о характеристиках областей, соответствующих связным множествам прецедентов, а также о взаимосвязях между ними.

Вторая часть доклада посвящена описанию системы, реализующей предложенный подход. Система предоставляет пользователю возможность построения произвольных ориентированных сетей из алгоритмов классификации и вычисления новых признаков, а также позволяет переходить от классификации прецедентов к классификации множеств прецедентов. Архитектура системы позволяет легко добавлять новые алгоритмы, что обеспечивает возможность ее применения к различным задачам из рассматриваемого класса.

#### **Литература**

1. Amin P. Charaniya, Roberto Manduchi, Roberto M., Suresh K. Lodha Supervised Parametric Classification of Aerial LiDAR Data // IEEE Conference on Computer Vision and Pattern Recognition Workshop; 2004; С. 25-32
2. Dragomir Anguelov, Ben Taskar, Vassil Chatalbashev, Daphne Koller, Dinkar Gupta, Jeremy Heitz, Andrew Ng Discriminative Learning of Markov Random Fields for Segmentation of 3D Scan Data // Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 2; 2005; С. 169 - 176
3. J. Reitberger, P. Krzystek, U. Stilla Analysis of full waveform LIDAR data for the classification of deciduous and coniferous trees // International Journal of Remote Sensing Volume 29, Issue 5 (March 2008); 2008; С. 1407-1431

### **Разрешимость задач нечёткой классификации элементов точечных конфигураций**

*Дорофеев Николай Юрьевич*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,*

*ВМиК, Москва, Россия*

*email: cmc.nick@gmail.com*

Рассматривается задача построения обучаемых алгоритмов классифика-

ции точек в плоских конфигурациях. Постановка задачи дана в [1]. Требуется каждой точке конфигурации на основании некоторой её окрестности сопоставить элемент из некоторого множества, называемого словарём разметки. В работе [2] были рассмотрены вопросы разрешимости задач выделения трендов. Указанные задачи были сведены к задаче классификации точек в плоских точечных конфигурациях. Там же были даны определения и получены критерии локальной разрешимости этих задач.

Полученные критерии опирались на понятие сдвиг-эквивалентности конфигураций и сдвиг-эквивалентности окрестностей, в котором конфигурации или окрестности считались эквивалентными, если они совпадали с точностью до сдвига. Неразрешимыми в этом случае считались задачи, в которых эквивалентным окрестностям соответствовали различные разметки. Однако, в случае незначительных различий конфигураций или окрестностей, которые могут возникать, например, как следствие погрешностей измерений или шумов в данных, отношение сдвиг эквивалентности будет потеряно. В этом случае задача становится формально разрешимой.

Одним из возможных путей решение данной проблемы является уменьшение точности измерения и хранения данных. Тогда, при удачно подобранной минимальной единице измерения, большая часть шумов не будет фиксироваться. Следовательно, мало отличающиеся конфигурации станут сдвиг-эквивалентными.

Другим выходом из обозначенной проблемы является переход от дискретного понимания поставленной задачи классификации к непрерывному. Введя меру близости окрестностей, степень различия двух окрестностей можно рассматривать как непрерывную величину. После введения меры близости для разметок становится возможным в качестве критерия разрешимости рассматривать соблюдение достаточно естественного требования сопоставлять близким окрестностям близкие метки. Алгоритм, удовлетворяющий такому требованию, будет аналогом липшиц-непрерывной функции, действующей из пространства окрестностей в пространство меток.

В докладе ставятся задачи перехода от дискретного понимания задач классификации элементов точечных конфигураций к непрерывному и поиска критериев разрешимости таких задач.

### Литература

1. Чехович Ю.В. Об обучаемых алгоритмах выделения трендов // Искусственный интеллект (научно-теоретический журнал НАН Украины). 2002. № 2. С. 298-305.
2. Рудаков К.В., Чехович Ю.В. Алгебраический подход к проблеме синтеза обучаемых алгоритмов выделения трендов // ДАН. М., 2003. Т. 388. № 1. С. 33-36.

# Условия корректности линейного замыкания модели АВО

*Иофина Галина Владимировна*

*Аспирант*

*МФТИ (ГУ), ФУПМ, Москва, Россия*

*email: giofina@gmail.com*

В данной работе развивается идея алгебраического подхода в задачах распознавания, предложенная Ю.И. Журавлёвым [4]. Ранее изучались линейные и алгебраические замыкания алгоритмов, получаемых путём изменения параметров. Для алгоритмов вычисления оценок (АВО) изменяющимися параметрами были веса объектов из обучающей выборки, веса признаков, опорные множества, нормировочные коэффициенты и т.д. Метрики на признаках были фиксированы и не являлись параметрами алгоритма. В данной работе фиксировались все параметры алгоритма кроме метрик на признаках, и были найдены критерии корректности линейного замыкания АВО для задач распознавания с непересекающимися и пересекающимися классами. Использовались обозначения, введённые в [2,3,5].

Для задач с непересекающимися классами была доказана

**Теорема 1.** Линейное замыкание  $L(B^*)$  задачи распознавания с непересекающимися классами при возможности выбора метрик на порядковых признаках произвольным образом корректно тогда и только тогда, когда выполнены первое и третье условия регулярности (множества эталонов каждого из классов попарно различны, и в контрольной выборке нет ни одной пары объектов, неразличимых относительно эталонов).

Для задач с пересекающимися классами с помощью элементарных свойств линейно зависимых систем векторов [1] было получено следующее условие корректности.

**Теорема 2.** Линейное замыкание  $L(B^*)$  задачи распознавания с пересекающимися классами при возможности выбора метрик на порядковых признаках произвольным образом корректно тогда и только тогда, когда выполнены первое и третье условия регулярности, и для каждого класса  $K_j$ ,  $j = 1, \dots, l$ , выполняется условие  $\text{Rg}(A^{K_j}) < i_{n_j}$  (здесь матрица  $A^{K_j}$  размера  $l \times i_{n_j}$  — объединение строк матрицы  $A$  принадлежности соответствующих объектов из  $K_j$  всем классам).

## Литература

1. Беклемишев Д.В. Курс аналитической геометрии и линейной алгебры. М., 1971.
2. Дьяконов А.Г. Алгебры над алгоритмами вычисления оценок: Учебное пособие, М, 2006.
3. Дьяконов А.Г. Критерии корректности алгебраических замыканий модели алгоритмов вычисления оценок // ДАН. 2008. Т. 420, № 6. С. 732–735.

4. Журавлёв Ю.И. Об алгебраическом подходе к решению задач распознавания или классификации. Проблемы кибернетики, выпуск 33. 1978. С. 5–68.
5. Матросов В.Л. Корректные алгебры ограниченной ёмкости над множеством регулярных задач: Дис. докт. физ.-матем. наук. М., 1985.

## **Применение генетических алгоритмов для выделения событий в информационном потоке.**

*Карпов Илья Андреевич*

*Студент*

*Национальный исследовательский ядерный университет «МИФИ»,*

*Факультет кибернетики, Москва, Россия*

*email: karilan@yandex.ru*

Сегодня трудно переоценить значимость анализа информации. Одной из возможных областей применения является определение актуальных тем, которые активно обсуждаются людьми в сети. Это могут быть почтовые сообщения, разговоры в чатах и форумах. Хотя активность вокруг какой-либо темы может быть легко обнаружена человеком, выделить изменение активности автоматически гораздо труднее.

Вот основные сложности при решении этой задачи:

- Периодические изменения и колебания активности.
- Модель должна быть универсальна для любой предметной области.
- Резкое повышение и спад активности могут произойти в любой момент и быть вызваны неизвестными системе факторами.
- Жесткие требования к производительности.

Идея подхода заключается в построении при помощи генетического алгоритма некоторой аппроксимирующей функции, описывающей “обычную” активность без учета случайных всплесков и последующего ее сравнения с реальными данными. В случае значительных расхождений с эталоном регистрируется всплеск активности.

При построении генетического алгоритма, в качестве популяции используется набор состояний вероятностного автомата [2]. Для получения функции приспособленности применялась формула вероятности изменения состояния для пуассоновского распределения, предложенная Клеинбергом в [1].

На рис. 1 представлен пример работы метода где приводятся выделение событий (выделены красным) для ключевого слова “june”. В качестве источника сообщений используется сервис микроблоггинга twitter.com. Построенная аппроксимирующая функция соответствует реальной зависимости на всем интервале кроме случаев появления случайных событий.

Главным преимуществом предложенного метода является низкое требование к вычислительным ресурсам. Сравнение результатов показало, что по

сравнению с методами регрессионного анализа предложенный метод дает схожие по точности результаты. При этом вычислительные затраты, особенно на больших массивах данных, значительно меньше.

### Литература

1. Kleinberg J.M. Bursty and hierarchical structure in streams // Data Min and Know Disc 2003.
2. Rabiner L.R. A tutorial on hidden Markov models and selected applications in speech recognition. 1990.

### Иллюстрации

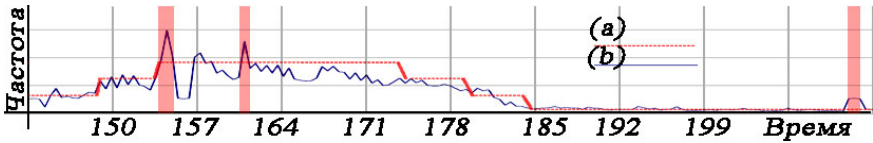


Рис. 1. График прогнозируемой функции (a) и реальной зависимости частоты вхождений (b) от времени.

## Разделение системы точек на подмножества с невырожденными матрицами попарных расстояний

*Карович Павел Алексеевич*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова*

*Факультет вычислительной математики и кибернетики, Москва,*

*Россия*

*email: mail@gmail.com*

Система  $q$  точек в пространстве  $\mathbb{R}^m$  с  $l_1$  метрикой называется 1-сингулярной, если матрица попарных расстояний для нее является вырожденной. Критерий 1-сингулярности был получен в работе [2].

В докладе мы расскажем о построении алгоритма разделения системы точек на подмножества без свойства 1-сингулярности. Будет показано, что задача сводится к построению алгоритма разделения множества векторов в конечномерном евклидовом пространстве на подмножества, каждое из которых является линейно-независимой системой векторов. Данная задача может быть успешно решена [1] с помощью теории матроидов [3]. Мы расскажем об алгоритме разделения и приведем оценку на минимальное число подсистем с невырожденными матрицами попарных расстояний, на которые может быть разбита произвольная система точек.

**Теорема.** Произвольная система  $S$  точек в  $\mathbb{R}^m$  может быть разбита на

$$\left[ \frac{\prod_{i=1}^m c_i}{\sum_{i=1}^m c_i + 1 - m} \right]$$

подмножеств с невырожденными матрицами попарных  $l_1$ -расстояний, где  $c_i$  - число значений, которые принимает  $i$ -я координата точек системы .

### Литература

1. Edmonds J. Matroid partition // Math. Decision Sciences. Proceedings 5th Summer Seminary Stanford. Part 1 (Lectures of Applied Mathematics 11). 1968. P. 335-345.
2. Reid L., Sun X. Distance matrices and ridge function interpolation // Canadian Journal of Mathematics. 1993. N 45. P. 1313-1323.
3. Schrijver A. Combinatorial optimization: polyhedra and efficiency. Berlin: Springer., 2003. 1. P. 651-761.

## Порождение и отбор признаков при восстановлении линейной регрессии

*Крымова Екатерина Александровна*

*Студент*

*Московский физико-технический институт, Факультет Управления и Прикладной Математики, Москва, Россия*  
*email: ekkrym@gmail.com*

Решение практических задач восстановления регрессионной зависимости требует рассмотрения большого числа признаков. Линейной модели недостаточно для построения модели удовлетворительного качества, поэтому происходит порождение новых признаков. При порождении признаков часто возникает проблема мультиколлинеарности. В данной работе для устранения мультиколлинеарности используется отбор признаков.

Процедура построения регрессионных моделей состоит из двух шагов [1]. На первом шаге, на основе свободных переменных - результатов измерений, порождается набор признаков. На втором шаге производится выбор признаков.

Предлагается следующий способ порождения признаков. Задано множество свободных переменных  $\Xi = \{\xi^u\}_{u=1}^U$ . Задано конечное множество функций

$G = \{g_v\}_{v=1}^V$ . Эти функции называются порождающими функциями (или примитивами). Обозначим  $a_i = g_v(\xi^u)$ , где индекс  $i = (v-1)U + u$ . Рассмотрим декартово произведение  $G \times \Xi$ , элементу  $(g_v, \xi^u)$  которого поставлена в соответствие суперпозиция  $g_v(\xi^u)$ , однозначно определяемая индексами  $v, u$ .

В качестве модели, описывающей отношение между зависимой переменной  $y$  и свободными переменными  $a_i$ , используется полином Колмогорова-Габора:

$$y = \beta_0 + \sum_{i=1}^{UV} \beta_i a_i + \sum_{i=1}^{UV} \sum_{j=1}^{UV} \beta_{ij} a_i a_j + \dots,$$

где вектор  $\beta$  - вектор параметров.

Переобозначим  $\beta = \{\beta_1, \beta_2, \dots, \beta_n\}$ . Запишем вышеприведенный ряд в виде линейной комбинации порожденных переменных,

$$y = \sum_{j=1}^n \beta_j x_j.$$

Переменные  $x_j$  поставлены в однозначное соответствие мономам полинома.

Требуется найти такое подмножество признаков  $\hat{\mathbf{x}} = \{x_i^j | j \in \mathcal{A}\}$ , заданное множеством индексов  $\mathcal{A} \subset N$ , для которого модель  $y = \sum_{j=1}^n \beta_j \hat{x}_j$ , где  $\beta$  - параметры модели, найденные по множеству признаков  $\hat{\mathbf{x}} = \{x_i^j | j \in \mathcal{A}\}$ , доставляет оптимальное значение заданному критерию качества.

В данной работе рассматриваются два алгоритмы отбора признаков на основе метода наименьших углов (LARS) [2] с возможностью удаления неинформативных признаков. Алгоритмы сравниваются с общеизвестными алгоритмами отбора признаков.

Сравнение алгоритмов выполнялось на исторических данных торгов опционом Brent Crude Oil [3]. Для каждого алгоритма вычислены значения различных критериев качества. Результаты сравнения алгоритмов проанализированы.

## Литература

1. Стрижов В. В. Методы индуктивного порождения регрессионных моделей // Сообщения по прикладной математике. Вычислительный центр РАН. 2008.
2. Efron B., Hastie T., Johnstone I., Tibshirani R. Least Angle Regression // The Annals of Statistics. 2004. Vol. 32, No. 2. Pp. 407–499.
3. Hull J. C. Options, Futures and Other Derivatives. Prentice Hall. 2000.

## Поиск взаимозаменяемых товаров

*Ломакин Василий Дмитриевич*

*Студент*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: v1-88@yandex.ru*

В настоящее время ассортименты магазинов крупных торговых сетей содержат десятки тысяч уникальных товарных единиц. Обнаружение в них

кластеров взаимозаменяемых товаров (субститутов) может иметь несколько практических применений. Во-первых, это позволяет решить задачу формирования оптимального ассортимента [1]. Не каждый магазин сети способен выставить на витрину весь ассортимент продукции. Использование как минимум одного товара из каждой взаимозаменяемой группы позволит максимально удовлетворить запросы покупателей. Во-вторых, учёт информации о субстигутах может повысить качество прогнозирования потребительского спроса. В-третьих, эта информация позволяет прогнозировать спрос на новые товары, только появившиеся на рынке, а также поведение покупателей при уходе товара.

Факт замещения с большой долей надёжности можно зафиксировать в моменты, когда по одному из товаров возникает дефицит, что вызывает незапланированное увеличение продаж других товаров группы. В данной работе исследуется возможность выявления субститутов на основе значимого повышения спроса на одни товары в периоды дефицита других.

Рассматривается следующая модель. Временные ряды продаж  $v_i(t)$ ,  $i = 1 \dots N$ ,  $t \in T$  представлены постоянными уровнями с добавлением гауссовского шума. Фиксируется ряд  $s$ , имеющий на отрезке  $T_0 \in T$  нулевые продажи (дефицит). Прогнозируемый объём продаж товара  $s$  на отрезке дефицита равен  $V$ . Принимается гипотеза о том, что объём спроса  $V$  полностью распределяется по другим товарам:  $\sum_{i=1}^N w_i = V$ , где  $w_i$  — доля товара  $i$  в замещении товара  $s$ . Задача состоит в вычислении долей  $w_i$ , значимо отличающихся от нуля. Предлагаются и сравниваются различные методы её решения. Согласно «наивному» методу доли  $w_i$  вычисляются как разность средних продаж товара  $v_i$  на отрезке дефицита  $T_0$  и средних продаж вне отрезка дефицита. Отрицательные и незначимые доли выделяются методами дисперсионного анализа [2] и обнуляются, оставшиеся нормируются на величину  $V$ . Во втором методе вместо нормировки настраивается уровень значимости так, чтобы сумма максимальных положительных долей товаров составила ровно  $V$ . В третьем методе для выделения товаров со значимыми положительными долями сравниваются распределения положительных и отрицательных долей.

Исследуются границы применимости предложенных методов на модельных данных. Определяются пороги общего размера ассортимента, уровня шума, длины отрезка дефицита, уровня продаж замещаемого товара, при превышении которых методы становятся неадекватными. Показано, что количество товаров в замещающей группе не влияет на качество работы алгоритма, а использование статистических тестов для выделения весов, близких к нулю, улучшает качество работы модели.

## Литература

1. Шрайбфедер Д. Эффективное управление запасами. 2-е изд. М., 2006.
2. Лапач С. Н., Чубенко А. В., Бабич П. Н. Статистика в науке и бизнесе. Киев: Морион, 2002.



# Экспериментальное исследование алгоритма отбора подмножеств взаимосвязанных признаков в задаче обучения распознаванию образов

*Маслакова Светлана Николаевна*

*Студент*

*Тульский государственный университет, Факультет кибернетики,  
Тула, Россия*

*email: svetlana.maslakova@gmail.com*

Проблема решения задачи обучения распознаванию образов в случае большого числа признаков, как правило, сводится к паре стандартных подходов – сокращению размерности за счет отбора наиболее информативных признаков и наложению на решающее правило априорных ограничений (регуляризация решающего правила распознавания). В работе исследуется способ, комбинирующий две эти методики. Рассматривается отбор информативных признаков в случае их упорядоченности в задаче обучения распознаванию образов. В ряде задач признаки – есть суть последовательных измерений вдоль оси некоторого аргумента, например, упорядоченные отсчеты какого-либо сигнала, компоненты спектра и т.п.

\*\*\*

Для исследования были выбраны следующие методы: метод потенциальных функций, и вытекающий из него метод опорных векторов (этот метод используется в качестве теоретической базы исследуемых алгоритмов); регуляризация решающего правила распознавания, учитывающая априорную информацию о взаимосвязи признаков [4, 5]; отбор информативных признаков, предложенный в работах [2, 3]; отбор подмножества признаков с учетом квадрата разности соседних весовых коэффициентов и отбор подмножества признаков с учетом модуля разности соседних весовых коэффициентов [1].

Ряд проведенных экспериментальных исследований как на модельных данных, так и на реальных показал, что введение регуляризации, опирающейся на поиск информативной подобласти признакового пространства, позволяет улучшить прогнозирующие свойства решающего правила по сравнению со стандартной процедурой метода опорных векторов. При этом процедура с применением штрафа в виде модуля точнее выделяет информативную подобласть признакового пространства, то есть дает меньший процент ошибки распознавания, по сравнению с процедурой с использованием квадратичного штрафа.

## Литература

1. Копылов А.В., Середин О.С., Приймак А.Ю., Моттль В.В. Отбор подмножеств взаимосвязанных признаков на основе параметрической процедуры динамического программирования // Всеросс. конф. ММРО-14 – М.: МАКС Пресс, 2009 – С. 136-140.
2. Mottl V.V., Seredin O.S., Krasotkina O.V., and Muchnik I.B. Fusing of potential functions in reconstructing dependences from empirical data In:

Doklady Mathematics, Vol. 71, No. 2, 2005, pp. 315–319. From Doklady Akademii Nauk, Vol. 401, No. 5, 2005, pp. 607–612.

3. Mottl V.V., Seredin O.S., Krasotkina O.V., and Muchnik I.B. Principles of multi-kernel data mining. In: P. Perner and A. Imiya (Eds.), Machine Learning and Data Mining in Pattern Recognition, Springer Verlag, LNAI 3587, 2005, pp. 52 – 61.
4. Seredin O.S., Dvoenko S.D., Krasotkina, O.V. and Mottl V.V., Machine Learning for Signal Recognition by the Criterion of Decision Rule Smoothness. Pattern Recognition and Image Analysis, Vol. 11, No. 1, 2001, pp. 87–90.
5. Seredin O., Mottl V. Regularization in Image Recognition: the Principle of Decision Rule Smoothing In: Proceedings of the Ninth International Conference Pattern Recognition and Information Processing, Minsk, Belarus, 2007. Vol.II., pp. 151-155.

### Слова благодарности

Выражаю благодарность своему научному руководителю Середину Олегу Сергеевичу.

### О метрическом аналоге метода главных компонент *Найденоев Никита Анатольевич*

*Студент*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: [naidyonov@gmail.com](mailto:naidyonov@gmail.com)*

В интеллектуальном анализе данных классическим является представление информации об объектах в виде признаковых описаний. Общеизвестны проблемы, которые возникают при плохом наборе признаков или при достаточно большом их количестве. Чтобы справиться с указанными проблемами, предложены различные подходы, в числе которых выделяют методы понижения размерности. Один из наиболее известных – это метод главных компонент (Principal components analysis, PCA). Классический PCA [1] формулируется как задача аппроксимации числовых признаковых описаний объектов, причем на признаки не накладываются никакие дополнительные ограничения.

В настоящее время всё больше прикладных и теоретических работ ориентировано на описание набора объектов не признаками, а попарными расстояниями между объектами. Такие описания называют метрическими. При этом возникают и требуют решения проблемы, аналогичные проблемам признаковых описаний [2]. В работе предложен и исследован аналог метода главных компонент для метрических описаний. На получаемое новое описание

малой размерности накладывается дополнительное ограничение в виде условия, что каждая новая компонента описания должна удовлетворять аксиомам полуметрики. Рассматриваемый метод назван метрическим методом главных компонент.

Доказана теорема о сохранении качества аппроксимации при добавлении метрических условий, т.е. дополнительные требования не снижают качества аппроксимации. Показано, что «привлекательные» свойства классического метода главных компонент сохраняются: главные компоненты можно вычислять частично и они годятся сразу для всех размерностей, главные компоненты ортогональны и можно удобно переходить от старых описаний к новым.

Проведены эксперименты на модельных и реальных метрических данных для анализа необходимости и роли свободного члена в методе главных компонент.

### Литература

1. T. Jolliffe Principal Component Analysis (2nd Ed.) // Springer-Verlag New York Inc. 2002
2. Майсурадзе А.И. О поиске оптимального коллективного слагаемого для набора метрических конфигураций // Искусственный интеллект , 2006, №2, с.146-150

### Слова благодарности

Автор выражает признательность к.ф.-м.н. Майсурадзе Арчилу Ивериевичу

## Инкрементные методы в коллаборативной фильтрации

*Полежаева Елена Андреевна*

*Студент*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,*

*Россия*

*email: lena\_polejaeva@mail.ru*

Методы коллаборативной фильтрации (CF) используются в рекомендующих системах и системах управления взаимоотношениями с клиентами (CRM) для автоматического формирования персональных предложений. Исходными данными является матрица  $Y$ , строки которой соответствуют  $n$  клиентам, столбцы —  $d$  объектам. Объектами могут быть товары, услуги, и т. п. Каждая заполненная ячейка матрицы содержит информацию об использовании данным клиентом данного объекта. Это может быть отметка о посещении, выставленный клиентом рейтинг, и т. д. Задача состоит в том, чтобы для произвольного клиента спрогнозировать оценки предпочтительности объектов по всем незаполненным ячейкам в строке матрицы  $Y$ . Предполагается,

что для этого выделяется множество клиентов со схожими предпочтениями (коллаборация).

Простые методы CF, основанные на поиске корреляций между клиентами (user-based) или объектами (item-based), неэффективны по времени и по памяти. Они требуют хранения всей матрицы  $Y$ . Этих недостатков лишены методы сингулярного разложения SVD, неотрицательного матричного разложения NNMF, позволяющие формировать сжатые описания (профили) клиентов и объектов. В современных динамических приложениях к методам CF предъявляется также требование инкрементности: метод должен эффективно пересчитывать хранимую информацию в случаях появления (1) нового клиента или объекта или (2) нового значения в ячейке матрицы  $Y$ .

В данной работе предлагается объединить оба типа инкрементности. При изменении одной ячейки матрицы  $Y$  применяются формулы инкрементного вычисления корреляции Пирсона (IPC) [2]. Разбиение формулы на части и хранение промежуточных величин повышает эффективность вычисления оценок сходства клиентов и объектов. При добавлении клиента или объекта применяется инкрементное сингулярное разложение ISVD [1], в результате которого получается новое SVD  $Y^{n \times d} = U^{n \times r} S^{r \times r} (R^T)^{r \times d}$ . Матрица  $US$ , содержащая профили клиентов, используется в формуле IPC. Пусть  $k$  — число клиентов, имеющих хотя бы с одним другим клиентом общий объект,  $n'$ :  $n' \ll n$  — число клиентов, с которыми данный клиент имеет хотя бы один общий объект,  $d'$ :  $d' \gg d$  — количество объектов, не оцененных данным клиентом, но оцененных по крайней мере одним из близких к нему клиентов. Объем хранимых данных  $O(n + kn' + (n + d)r)$ . Добавление клиента требует  $O((nr + d)r)$  операций, добавление объекта —  $O((nr + d)nr)$  операций, модификация ячейки матрицы  $Y$  —  $O(n'd')$  операций. Таким образом, предложенный инкрементный метод CF достаточно эффективен для использования в современных приложениях.

## Литература

1. M.Brand. Fast Low-rank modifications of the thin singular value decomposition // Linear Algebra and Its Applications, Vol. 415, Issue 1, Pp. 20–30, 2006.
2. M.Papagelis. Incremental collaborative filtering for highly-scalable recommendation algorithms. // ISMIS, Springer, Vol. 3488, Pp. 553–561, 2005.

## Слова благодарности

Исследования поддержаны грантом РФФИ 10-07-00609.

# Поиск похожих пользователей в социальных сетях методами коллаборативной фильтрации

*Пустовойтов Никита Юрьевич*

*Соискатель*

*Московский Физико-Технический Институт (МФТИ), ФУПМ, Москва,  
Россия*

*email: flashnik86@gmail.com*

В настоящее время в Интернете наблюдается бурный рост социальных сетей. В данной работе делается попытка извлечь ценную информацию из отношений между пользователями и их описаниями. При этом учитывались только списки друзей (friend list) и ключевые слова и выражения, которыми пользователи описали свои интересы в социальной сети. Далее любое ключевое слово или целое выражение будем называть просто интересом.

Для наглядности удобно представлять исходные данные в виде матрицы, как это часто делают в задачах коллаборативной фильтрации. Пусть задано множество пользователей  $\mathbf{U}$  ( $|\mathbf{U}| = m$ ) и множество интересов  $\mathbf{R}$  ( $|\mathbf{R}| = n$ ). Заданы матрица интересов  $A_{m \times n}$  и матрица дружбы пользователей  $F_{m \times m}$ . Строки матриц  $A$  и  $F$  соответствуют пользователям, столбцы матрицы  $A$  — интересам, столбцы матрицы  $F$  — пользователям. Элемент матрицы  $A_{ij} = 1$  ( $F_{ij} = 1$ ), если пользователь  $U_i$  указал интерес  $R_j$  (добавил  $U_j$  в друзья), и 0 в противном случае. Матрица дружбы  $F$  не является симметричной. При этом обе матрицы являются сильно разреженными.

Требуется построить рекомендательный алгоритм, заполняющий пустые значения в матрицах  $A$  и  $F$  (или одной из них) таким образом, чтобы доставить выбранному функционалу качества наибольшее значение.

Так как содержательно важным является только значение элемента матрицы, равное 1 (в большинстве случаев 0 означает что пользователь «не знает»), то алгоритм должен восстанавливать только единичные значения в этих матрицах. Восстановление  $A_{ij} = 1$  или  $F_{ij} = 1$  означает, что пользователю  $U_i$  рекомендуется, соответственно, интерес  $R_j$  или пользователь  $U_j$  для добавления в друзья.

Будем считать, что любой объект рассматриваемой задачи можно описать вектором вещественных чисел  $P$  заранее фиксированной размерности  $T$ , для компонент которого выполняются условия  $P_t \geq 0 \forall t \in T$  и  $\sum_{t=1}^T P_t = 1$ . Такой вектор скрытых переменных (latent variables) будем называть скрытым профилем объекта. Компонента скрытого профиля интерпретируется как оценка близости объекта некоторой теме.

При известных скрытых профилях объектов для заполнения искомым матриц достаточно найти расстояние между векторами и сравнить его с некоторым порогом. Таким образом, задача сводится к выявлению скрытых профилей.

В данной работе предлагается вероятностная модель данных, описывающая связь скрытых профилей пользователей и интересов. Это позволило для восстановления скрытых профилей использовать EM-алгоритм.

Обучение EM-алгоритма для настройки профилей также применялось в [1] для анализа поисковых логов. Предложенный в нашей работе алгоритм на рассматриваемой нами задаче показал более высокую скорость сходимости, чем наша реализация [1].

### Литература

1. Leksin V. A. Symmetrization and overfitting in probabilistic latent semantic analysis // Pattern Recognition and Image Analysis. - 2009. - Vol. 19, No. 4 - P. 565-574.

## Оценки вероятности переобучения, зависящие от графа расслоения и связности семейства алгоритмов

*Решетняк Илья Михайлович*

*Студент*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: ananace@gmail.com*

При решении задач классификации часто возникает проблема переобучения, когда алгоритм, выбранный из заданного семейства алгоритмов  $A$  по критерию минимума числа ошибок на обучающей выборке, существенно чаще ошибается на независимых контрольных данных. Классические оценки вероятности переобучения зависят только от длины выборки и сложности семейства  $A$ . Сложность определяется как число различных бинарных векторов  $\vec{a} = (b_{a1}, \dots, b_{aL})$ , порождаемых семейством  $A$  на выборке  $X^L = \{x_1, \dots, x_L\}$ , где  $b_{ai} = 1$  означает, что алгоритм  $a$  допускает ошибку на объекте  $x_i$ ,  $L$  — суммарная длина обучающей и контрольной выборки. Сложностные оценки пессимистично завышены. Для получения точных оценок необходимо учитывать эффекты расслоения и связности в семействе алгоритмов [1]. Расслоением называется разбиение  $A = A_1 \cup \dots \cup A_L$ , где слой  $A_m$  — это множество алгоритмов, допускающих  $m$  ошибок на выборке  $X^L$ . Эффект связности возникает, когда для каждого алгоритма  $a$  в семействе находятся такие алгоритмы  $a'$ , что векторы ошибок  $\vec{a}, \vec{a}'$  отличаются только на одном объекте. На практике это происходит, когда  $A$  представляет собой параметрическое семейство функций, непрерывных по параметрам. Семейство алгоритмов удобно представлять в виде ориентированного многодольного графа, вершины которого соответствуют векторам  $\vec{a}$ , рёбрами соединяются векторы, отличающиеся на одном объекте, от меньшего числа ошибок к большему, а доли соответствуют слоям.

Получена оценка вероятности переобучения, зависящая от совокупности величин  $\Delta_{mq}$ , равных числу алгоритмов в слое  $A_m$ , связанных рёбрами с  $q$  алгоритмами в слое  $m + 1$ . Аналогичная оценка получена также в [1], но при сильном дополнительном предположении, что в семействе существует алгоритм, не допускающий ошибок на  $X^L$ . В данной работе делаются более сла-

бые предположения — либо что граф семейства алгоритмов имеет единственный исток, либо что в семействе всегда можно найти алгоритм, не допускающий ошибок на обучающей выборке. В численном эксперименте показано, что полученные оценки значительно лучше классических и довольно близки к несмещённой эмпирической оценке, полученной методом Монте-Карло. В экспериментах с линейными алгоритмами классификации и методом ближайших соседей оказалось, что двумерный профиль расслоения и связности  $\Delta_{mq}$  достаточно точно приближается произведением одномерных профилей  $\Delta_{mq} \approx \lambda_q |A_m|$ , где  $\lambda_q$  — доля вершин графа, связанных рёбрами с  $q$  алгоритмами в следующем слое.

### Литература

1. Воронцов К. В., Рудаков К. В., Чехович Ю. В. О теоретико-множественных ограничениях и комбинаторной теории переобучения для алгоритмов классификации // Труды МФТИ, Т. 1, № 4 — 2009.

## Построение регрессионных моделей прогноза временных рядов

*Сологуб Роман Аркадьевич*

*Студент*

*Московский физико-технический институт, Факультет инноваций и  
высоких технологий, Москва, Россия*

*email: roman.sologub@yahoo.com*

Исследовано поведение акций и фьючерсов ОАО "Газпром". Цель исследования заключается в поиске оптимального алгоритма нахождения точек неэффективности рынка. Алгоритм используется для определения момента покупки/продажи инструмента. При этом рынок рассматривается как набор временных рядов. Гипотеза эффективности рынка предполагает, что гарантированно извлекать прибыль из движения цены инструмента невозможно. Тем не менее, на реальном рынке существуют инструменты, несоответствующие данной гипотезе, — например инструмент, состоящий из купленного фьючерса и проданной акции. В дальнейшем такой пакет будет называться спредом. Искомый алгоритм должен определять оптимальные моменты для покупки и продажи спреда. Такие моменты определяются с помощью прогноза значений спреда на некотором временном промежутке. Процедура построения регрессионных моделей прогнозирования состоит из двух шагов [1]. На первом шаге, на основе свободных переменных — исторических данных, порождается набор признаков. На втором шаге с помощью этих признаков порождается обобщенная линейная регрессионная модель. Для решения задачи выбирается исторический временной интервал, на котором задана регрессионная выборка — множество пар  $D = \{(\mathbf{x}_n, y_n)\}_{n=1}^N$ , в котором  $\mathbf{x} \in \mathbb{R}^P$  — набор временных рядов и временной ряд  $y \in \mathbb{R}^1$  — зависимая переменная. В контексте рассматриваемой задачи наборы временных рядов  $\mathbf{x}_n, y_n$  содержат значения цен на выбранном интервале времени. Предлагается следующий способ

порождения признаков. Задано конечное множество порождающих функций  $G = \{g|g : \mathbb{R} \times \dots \times \mathbb{R} \rightarrow \mathbb{R}\}$ . Функция  $g = g(\mathbf{b}, \cdot, \cdot, \dots, \cdot)$  — гладкая параметрическая. Первый аргумент функции — вектор параметров, последующие аргументы — функции свободных переменных. Множество  $G$  индуктивно определяет набор допустимых суперпозиций  $F = \{f_i\}$ . Суперпозиция  $f$  задает параметрическую регрессионную модель  $f = f(\mathbf{w}, \mathbf{x})$ . Полученные модели  $f_i$  сравниваются в соответствии с принципами связанного Байесовского вывода. Критерием выбора оптимальной модели является её правдоподобие, рассчитываемое как плотность вероятности появления текущего набора данных при использовании данной модели  $p(D|f_i)$ . Эта плотность рассчитывается как произведение  $p(D|\mathbf{w}, f_i)p(\mathbf{w}|f_i)$ . Модели настраиваются в соответствии с критерием максимизации результата ретроспективного прогноза по периоду  $T$ :  $\mathbf{w} = \arg \max(Q(f(\mathbf{w}, \mathbf{x})))$ , где  $Q(f) = \sum_{t_i \in T} \{\text{sign}[f(\mathbf{w}, t_i)]y(t_i)\}$ . Функционал  $Q$  является суммарной прибылью от операций, а  $t_i$  определяется моделью  $f$ .

## Точная оценка вероятности переобучения для одного специального семейства алгоритмов

*Толстихин Илья Олегович*

*Студент*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,*

*Россия*

*email: xjucex@gmail.com*

С явлением переобучения часто приходится сталкиваться при решении задач классификации, регрессии, прогнозирования. Оно заключается в том, что алгоритм, найденный по критерию минимума частоты ошибок на обучающей выборке, может значительно чаще ошибаться на независимой контрольной выборке.

Пусть задано множество объектов  $\mathbb{X} = \{x_1, \dots, x_L\}$  и множество алгоритмов  $A = \{a_1, \dots, a_D\}$ . *Матрицей ошибок* называется бинарная  $L \times D$ -матрица,  $(i, j)$ -й элемент которой равен 1, если алгоритм  $a_j$  допускает ошибку на объекте  $x_i$ . *Методом обучения*  $\mu$  называется отображение, которое каждой обучающей выборке  $X \subseteq \mathbb{X}$  ставит в соответствие некоторый алгоритм  $a = \mu X$  из  $A$ . Обозначим через  $\nu(a, X)$  частоту (долю) ошибок алгоритма  $a$  на выборке  $X$ . Будем полагать, что все разбиения множества  $\mathbb{X}$  на наблюдаемую обучающую выборку  $X$  и скрытую контрольную выборку  $\bar{X} = \mathbb{X} \setminus X$  равновероятны. Задача состоит в том, чтобы найти вероятность переобучения  $Q_\varepsilon = \mathbb{P}[\nu(\mu X, \bar{X}) - \nu(\mu X, X) \geq \varepsilon]$  при произвольном  $\varepsilon \in [0, 1]$ .

Теория Вапника–Червоненкиса (VC) даёт для  $Q_\varepsilon$  завышенные оценки «худшего случая», зависящие только от размеров матрицы ошибок. В [1] показано, что для получения точных оценок  $Q_\varepsilon$  необходимо учитывать свойства расслоения и сходства множества вектор-столбцов матрицы ошибок. Эффект *расслоения* связан с тем, что алгоритмы с малой частотой ошибок  $\nu(a, \mathbb{X})$



имеют существенно большую вероятность  $P[\mu X = a]$  оказаться результатом обучения, и если таких алгоритмов немного, то  $Q_\varepsilon$  заметно понижается. Эффект *сходства* связан с тем, что наличие вектор-столбцов, близких в смысле метрики Хэмминга, также снижает  $Q_\varepsilon$ . Комбинаторный подход [1] позволяет учитывать оба эффекта и получать точные оценки  $Q_\varepsilon$ .

Чтобы определить, насколько существенен эффект сходства при отсутствии эффекта расслоения, в данной работе вводится модельное семейство, определяемое как множество вектор-столбцов, образованное пересечением  $m$ -го слоя булева куба  $\{0, 1\}^L$  с хэмминговым шаром радиуса  $r$  с центром в некотором элементе  $m$ -го слоя. С помощью теоретико-группового метода, предложенного в [2] для симметричных семейств алгоритмов, получена точная оценка вероятности переобучения для метода обучения  $\mu$ , минимизирующего число ошибок на обучающей выборке. Вычисления показывают, что полученная оценка на порядки точнее классической VC-оценки.

### Литература

1. Воронцов К. В. Точные оценки вероятности переобучения // Доклады РАН, 2009. — Т. 429, № 1. — С. 15–18.
2. Фрей А. И. Точные оценки вероятности переобучения для симметричных семейств алгоритмов // Всеросс. конф. ММРО-14. — М.: МАКС Пресс, 2009. — С. 66–69.

## Метод извлечения оценочных слов из текстов

*Четвёркин Илья Игоревич*

*Студент*

*Московский государственный университет имени М. В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: ilia2010@yandex.ru*

В настоящее время на страницах сети Интернет можно найти множество отзывов о тех или иных товарах, фильмах, книгах и т.п. Эти отзывы содержат много полезной информации, поэтому эту информацию важно автоматически извлекать и представлять пользователям [2].

Мнения пользователей о продукте часто выражаются посредством оценочных слов и выражений, которые несут в себе некоторую положительную или отрицательную оценку. Поэтому важным фактором качественного извлечения мнений о той или иной сущности является знание оценочных слов и выражений, которые используются в данной области. Проблема состоит в том, что невозможно заранее собрать список оценочных выражений, которые будут применимы для всех предметных областей, поскольку некоторые оценочные выражения употребляются только в конкретных предметных областях, другие являются оценочными в одной области и не являются оценочными в другой.

В работе рассматриваются методы автоматического получения оценочных слов на основе нескольких корпусов текстов, которые можно автоматически построить для многих предметных областей, а именно, корпуса отзывов о сущности с вручную проставленными потребителями оценками, корпуса нейтральных описаний сущностей и нейтрального контрастного корпуса, например, составленного из потока общезначимых новостей. Из указанных корпусов извлекаются списки слов, упорядоченные по значению различных признаков, оценивается качество этих списков относительно содержания в них оценочных слов по двум разным характеристикам, и исследуются способы комбинирования этих списков для получения лучшего качественного состава по оценочным словам. Всего было предложено 17 признаков слов, вычисленных на основе извлеченных корпусов.

Для оценки качества получаемых списков слов были адаптированы метрики оценки качества сортировки списков, применяемых для информационного поиска [1]: точность и средняя точность (MAP), которая является одной из основных мер, применяемых для оценки качества упорядочения в списках (например, качества выдачи при поиске по запросу).

Эксперименты проводились в предметной области отзывов о фильмах.

Извлечение и комбинирование слов производилось отдельно по прилагательным и не прилагательным словам. Такое разделение связано с тем, что наиболее часто в качестве оценочных слов применяются именно прилагательные. В результате комбинирования признаков удалось получить для тысячи прилагательных наилучшее значение средней точности - 81.5%, точности - 71.9 %.

В результате мы показали, что путем извлечения нескольких корпусов заданной предметной области и вычисления с их помощью нескольких характеристик слов можно автоматически получить достаточно качественные списки оценочных слов.

## Литература

1. Агеев М., Кураленок И., Некрестьянов И. // Петрозаводск: Российский семинар по Оценке Методов Информационного Поиска (РОМИП 2009), 2009.
2. Liu B. Opinion Mining and Summarization // WWW-2008, Beijing, 2008.

## Слова благодарности

Хочу выразить особую благодарность своему научному руководителю Лукашевич Наталье Валентиновне, за помощь, оказанную в написании работы.

# Самообучаемая система коррекции L<sup>A</sup>T<sub>E</sub>X-документов

*Чувилин Кирилл Владимирович*

*Студент (бакалавр)*

*email:*

Многие научные журналы и конференции принимают статьи в формате L<sup>A</sup>T<sub>E</sub>X. Как правило, авторские тексты содержат значительное количество типографических ошибок, которые исправляются корректорами вручную, что требует много времени и внимания. Предлагаемая система предназначена для автоматизации корректуры и существенного уменьшения времени ручной работы. В её основе лежит наблюдение, что большинство ошибок являются типовыми, и их исправление может быть формализовано с помощью системы правил макроподстановок (макросов), в простейшем случае заменяющих текст по шаблону [2].

Начальная система макросов формируется опытными корректорами. При обработке документа применяются все макросы, пока это возможно. В случае применимости макроса у пользователя запрашивается подтверждение на выполнение замены. Если пользователь соглашается со всеми заменами, то обработка происходит максимально быстро. Если пользователь не согласен с заменой, то он либо исправляет текст документа, либо изменяет макрос. Все исправления текста считаются отрицательными прецедентами, все согласия — положительными, и в совокупности они образуют обучающую выборку для данного макроса. При изменении макроса все ранее обработанные документы проверяются заново, чтобы гарантированно выполнялось требование «ни один из макросов не применим ни к одному из уже обработанных документов». Накопленная обучающая выборка используется для самообучения системы. При достаточном числе отрицательных примеров производится синтез новых макросов, уточняющих имеющийся так, чтобы отрицательные примеры устранились. Для каждого синтезированного макроса вычисляются показатели качества: число его применений, число устранённых ошибок, число новых ошибок при повторной проверке предыдущих документов.

Документы L<sup>A</sup>T<sub>E</sub>X представляются в виде дерева. Узлами дерева являются: команды, их аргументы, формулы, абзацы, слова, а терминальными вершинами — символы, отступы и промежутки. Такое представление позволяет в макросах учитывать не только последовательность символов, но и контекст. Для синтеза новых макросов применяются методы поиска логических закономерностей [1].

## Литература

1. Воронцов К.В. Методы машинного обучения, основанные на индукции правил // 17-я межд. конф. по понятийным структурам ICCS-17, семинар «Знания и онтологии ELSEWHERE 2009», Москва — С. 57–71.
2. Фридл Д. Регулярные выражения, 3-е издание. — Пер. с англ. СПб.: Символ-Плюс, 2008.

# Моделирование поведения группы агентов с помощью взаимосвязанных скрытых марковских моделей

*Юданов Анатолий Александрович*

*Студент*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики,  
Черноголовка, Россия*

*email: ХАРКОННЕН@gmail.com*

Моделирование сложных систем приобретает всё большую популярность в связи с активным ростом вычислительных мощностей и развитием новых подходов к моделированию. Статистические и нейросетевые модели применяются для распознавания речи, анализа статических и динамических изображений, для анализа поведения людей по данным видеокамер. Вид используемой модели зависит от предположений о внутренней структуре системы.

В работе рассматривается задача моделирования систем, представимых в виде группы равноправных взаимодействующих во времени агентов. Для таких систем предлагается строить модель в виде множества СММ с дополнительными связями специальной структуры. Модель подразумевает двухуровневое поведение каждого агента. Верхний уровень отражает реальные данные, полученные от агента, нижний уровень - стратегии его поведения. Стратегии агента зависят от стратегий всех агентов системы, поведение агента на верхнем уровне зависит только от его стратегий. Для предложенной модели выведены расширенные алгоритмы Баума-Уэлша и Витерби. Алгоритмы реализованы в виде программной системы ССММ.

Предложенный метод моделирования был применен для игр класса многопользовательский бескоалиционный шутер от первого лица. Агентами в играх этого класса являются люди-игроки и специализированные программы. Агенты взаимодействуют в трехмерной виртуальной среде с помощью ограниченного набора действий. Действия каждого агента ведут его к достижению некоторой цели, одновременное достижение целей несколькими агентами невозможно. Входом для системы ССММ является событийное поведение агентов. Для получения входных данных были написаны программы извлечения внутренней информации из игры. В режиме обучения ССММ настраивает матрицы переходов для состояний первого и второго уровня по обучающей выборке. В режиме анализа ССММ строит последовательность состояний нижнего уровня, разделяя последовательность событий верхнего уровня на тактические группы, имеющие осмысленную интерпретацию в терминах игры. Апробация системы ССММ проводилась на выборке из 12000 событий из реальной дуэльной игры, объём обучающей выборки 2000 элементов. В результате исходная выборка была размечена тремя тактическими группами: атака, поиск, бездействие после смерти.

Для моделирования поведения систем взаимодействующих агентов теоретически подходит дискретная СММ с векторными событиями. Однако коли-

чество элементов матрицы такой модели экспоненциально растёт от структурных параметров системы. Разработанная модель требует линейное число элементов матриц и, как показали проведенные тесты, лучше выделяет тактические элементы поведения агентов.

### Литература

1. Jun Yin, Yan Meng: Abnormal Behavior Recognition Using Self-Adaptive Hidden Markov Models // Volume 5627/2009, pp.337-346
2. Kosmopoulos D., Antonakaki P., Valasoulis K., Kesidis A., Perantonis S.: Human Behavior Classification Using Multiple Views // Volume 5138/2008, pp.123-134
3. Bishop C. M.: Pattern recognition & machine learning // ISBN: 978-0-387-31073-2, 2006

## Применение генетического программирования для обучения решающих деревьев в контексте процедуры бустинга.

*Янгель Борис Константинович*

*Студент*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: hrOnix@acm.org*

Бустинг [3] — один из методов, широко используемых в машинном обучении. Основная идея бустинга состоит в том, что неизвестная зависимость  $y(\vec{x})$  может быть приближена линейной комбинацией базовых решающих правил

$$\tilde{y}_T(\vec{x}) = \sum_{t=1}^T h(\vec{x}; \vec{\alpha}_t).$$

Обучение линейной комбинации выполняется жадным итерационным процессом. В качестве базовых решающих правил  $h(\vec{x}; \vec{\alpha})$  обычно используются решающие деревья небольшой высоты и пороговые классификаторы, как их частный случай.

Бустинг широко применяется в алгоритмах распознавания объектов на изображениях [4]. Он, однако, имеет существенный недостаток: время тренировки классификатора изображений методом бустинга может быть очень большим в силу высокой размерности признакового пространства на основе, например, признаков Хаара. Для бустинга пороговых классификаторов проблема с большим временем обучения была успешно решена в ряде работ (смотрите, например, [1], [2] и [5]). Тем не менее, предлагаемые в этих работах методы плохо применимы к решающим деревьям, использование которых в

качестве базовых решающих правил может существенно улучшить обобщающую способность классификатора.

В данной работе предлагается способ ускорить обучение линейной комбинации решающих деревьев методом бустинга. Для этого предлагается использовать комбинированную оптимизационную процедуру, выбирающую структуру решающего дерева при помощи эволюционного алгоритма, а остальные его параметры — путем точного решения соответствующих оптимизационных задач. Операторы мутации и скрещивания для эволюционного алгоритма предлагается позаимствовать из области генетического программирования, широко известной методики оптимизации древовидных структур.

Реализация предлагаемого метода свободно доступна по адресу <http://code.google.com/p/evoboost>. Результаты экспериментов с ней подтверждают высокую эффективность предлагаемого метода в задачах обучения классификаторов изображений, как с точки зрения скорости обучения, так и с точки зрения обобщающей способности результирующих решающих правил.

### Литература

1. Янгель Б.К. Ускорение бустинга параметрических классификаторов с использованием генетических алгоритмов // Сборник докладов 14й всероссийской конференции "Математические методы распознавания образов". М., 2009. С. 208-211.
2. M.-T. Pham, T.-J. Chain. Fast training and selection of Haar features using statistics in boosting-based face detection // IEEE 11th International Conference on Computer Vision. 2007.
3. R. Schapire. The boosting approach to machine learning: An overview // MSRI Workshop on Nonlinear Estimation and Classification. 2002.
4. P. Viola, M. Jones. Robust Real-time Object Detection // International Journal of Computer Vision. 2001.
5. J. Wu, C. Brubaker, M. Mullin, J. Rehg. Fast asymmetric learning for cascade face detection // IEEE Transactions on Pattern Analysis and Machine Intelligence volume 30. 2008.

# Парадигмы программирования и интеллектуальная обработка данных

## Онтологический подход для репрезентации интенциональной парадигмы языка MSC

*Вдовенко Татьяна Владимировна*

*Студент*

*Киевский Национальный Университет имени Тараса Шевченко,*

*Факультет кибернетики, Киев, Украина*

*email: htam7@mail.ru*

Большинство языков программирования созданы на подобии естественного языка, такой подход включает в себе удобство для практического применения. Однако в языках программирования, в отличие от естественного языка, есть только экстенциональные, ориентированы на идентификацию разных значений, стороны, в то время как интенциональные стороны, ориентированные на идентификацию разных смыслов, практически опускаются. Интенциональность служит для заложения в программу содержательной стороны алгоритма, структуры данных, что могло бы оказать существенную помощь для тестирования, представления программ в виде черных и белых ящиков[2]. При исследовании экстенциональности в мышлении (линия исследований Фреге – Карнапа[1]) как важную особенность можно отметить то, что способность к экстенциональности есть апостериорной (определяющей опытом) и возрастающей в процессе обучения; а к интенциональности априорной (не зависимой от опыта)[1], а в программных системах экстенциональность сразу в полном объеме закладывается разработчиком, а интенциональность практически не отражается.

В ходе работы на основе грамматики языка MSC были выделены базовые экстенциональные элементы сценарного программирования на языке MSC, на их основе были спрогнозированы интенционалы по методу[3]. Построена иерархическая градация этих элементов, от наиболее простых – к более сложным. Аналогичная градация была проведена для категорий данных, построена структура категориального определения данных от более простых к более сложным. Связующим звеном от неделимых базовых частей к слож-

ным структурам выступали семантические связи – траектории семантической взаимосвязи, последние также были типизированы по основным категориям. На основе выделенной информации была построена онтология для язык MSC. При обращении к онтологии идет прогноз возможных вариантов развития событий (интенционалы), и определения от конкретных концептов их представления на разных уровнях абстракции (локального и глобально-го экстенционалов). Способность обучения на новых сущностях заложена в онтологии: если встречается какая-то сущность, которая не даёт семантическую траекторию (интенционал), она входит в онтологию без семантических связей, впоследствии происходит ее обучение, в ходе которого формируются связи. Реализована возможность разного рода абстракций от несущественных концептов с целью более конкретного определения интенционала

Данный подход позволяет выделить определенные семантические траектории каждого концепта, объекта программы, что делает программирование проще, а способность к обучению адаптирует онтологию на специфику конкретной программы. Полученная онтология может быть применима в интерпретаторе языка MSC

### Литература

1. Карнап Р. Значение и необходимость. /Русский перевод. Под ред Д.А. Бачвара, М., 2007.
2. Никитченко Н. С. Интенциональные аспекты понятия программа. // Проблемы программирования. 2001. No 3-4. С. 5–13.
3. Gamut L. T. F. Logic, Language, and Meaning. Intensional Logic and Logical Grammar. London, 1991. V.2

### Слова благодарности

Выражаю благодарность доктору ф-мат. наук проф. Никитченко Николаю Степановичу за научное руководство

### Разработка системы интеллектуального анализа данных для оптимизации процесса электронного документооборота высшего учебного заведения

*Гладких Николай Александрович*

*Аспирант*

*Курский государственный университет, Факультет информатики и  
вычислительной техники, Курск, Россия*

*email: gladkih@kursknet.ru*

Комплексное обеспечение высшего учебного заведения средствами централизованного управления данными позволяет снизить временные и ресурсные затраты, повысить уровень качества делопроизводства [3]. Эффективность



процесса автоматизированного документооборота в рассматриваемой предметной области может быть в значительной степени снижена под действием ряда факторов: возникновение в процессе эксплуатации нештатных ситуаций, непредусмотренных на этапе проектирования системы электронного документооборота (СЭД); ошибки конечного пользователя (оператора, администратора системы) и др. Последствиями таких ошибок являются потеря или искажение данных, рассогласования и противоречия в работе различных структур вуза.

Это обуславливает актуальность разработки системы интеллектуального анализа данных (СИАД), задачами которой являются извлечение, накопление и обработка знаний, сценариев, моделей поведения информационной системы. Среди задач СИАД выделяются: анализ структуры баз данных, запросов, журналов транзакций, поиск закономерностей и взаимосвязей между данными, построение продукционных моделей, эмпирических моделей, деревьев решений и семантических сетей [2]. Предлагаемая нами методика основана на использовании регрессионного, дисперсионного, анализа данных, нейросетевых, генетических моделей [1]. Реализация СИАД предполагает разработку программного обеспечения – набора библиотек утилит, расширяющих функциональные возможности СЭД.

Данная СИАД включает в себя следующий набор функций: 1) построение сценариев поведения СЭД, основанное на опыте ее использования; 2) анализ кризисных ситуаций СЭД, обеспечивающий нахождение оптимальных путей их разрешения; 3) реализация функций сбора, хранения и систематизации статистических данных СЭД; 4) снижение рисков потерь или искажения данных; 5) обеспечение разработчиков ценной информацией о поведении СЭД, что позволит облегчить ее дальнейшую оптимизацию и поддержку. Внедрение предлагаемой методики связано с решением таких проблем, как: минимизация информационных рисков; согласование порядка эксплуатации СИАД с законодательством РФ; приведение программного и аппаратного обеспечения в соответствие с возникающими в процессе функционирования СИАД характером задач и уровнем нагрузки.

В докладе рассматриваются вопросы, связанные с построением СИАД в целях оптимизации процесса электронного документооборота образовательного учреждения. В качестве направлений дальнейшей работы предлагаются: адаптация и внедрение разработанной СИАД в другие корпоративные информационные системы (за пределами вуза); проектирование и реализация общеузовской базы знаний.

### Литература

1. Artificial Neural Networks in Real-Life Applications / Rabunal J. R., Dorrado J. (Eds.). Hershey – London – Melbourne – Singapore: Idea Group Publishing, 2006.
2. Data Mining: Opportunities and Challenges / Wang J. (Ed.). Hershey – London – Melbourne – Singapore – Beijing: Idea Group Publishing, 2003.

3. <http://www.ecm-journal.ru>.

### Слова благодарности

Автор выражает признательность научному руководителю, профессору, д. ф.-м. н. Добрице В. П. за помощь в работе над проектом.

### Поиск фразеологизмов в тексте

*Грибанов Д.В., Смышчѣк Е.М.*

*Студент*

*Нижегородский государственный университет имени Н.И.  
Лобачевского, Факультет вычислительной математики и  
кибернетики, Нижний Новгород, Россия  
email: dimitry.gribanov@gmail.com*

#### Поиск фразеологизмов в тексте

Нижегородским коллективом разработчиков ведется разработка системы способной распознавать фразеологизмы в тексте. Для повышения точности нахождения предполагается использовать синтаксические связи между словами в фразеологизме и словами в предложении. Для получения синтаксических связей строится дерево синтаксических связей предложения. Для построения дерева используется Stanford Parser[4,5]. А также предполагается функция приближенного нахождения фразеологизмов в тексте, которая предупреждает пользователя о словосочетаниях частично совпадающих с фразеологизмом по словам и синтаксическим связям.

\*\*\*

В работе рассмотрена сама структура фразеологизмов, и введено понятие, по-которому планируется отличать фразеологизм от обычного словосочетания и от другого фразеологизма. А именно под фразеологизм понимается набор его неделимых частей, между которыми есть синтаксические связи.

Разработаны методы основанные на алгоритме Ахо-Корасика[1,2] и алгоритме на суффиксном дереве[1,3]. Фактически они обобщают эти методы на понятие фразеологизма введенное выше. Над базой фразеологизмов при этом проводится пред-обработка связанная с одним из алгоритмов Ахо-Корасика или алгоритмом на суффиксном дереве, количество дополнительной памяти для пред-обработанной базы  $O(b+s)$ , где  $b$ -количество слов в базе, а  $s$  – количество символов в ней.

Приведенные методы выполняют задачу за время линейно зависящее только от длины предложения и независящую от размера базы фразеологизмов, что делает эти алгоритмы чрезвычайно эффективными.

Дополнительные преимущества данных методов:

1) Способность находить частичные включения фразеологизмов. Например, во фразе «биться головой о кирпичную стену» слово «кирпичную» можно выкинуть без утери смысла фразы. Точное частичное обнаружение фразеологизмов возможно благодаря анализу синтаксических связей.

- 2)Эффективность не только благодаря оценкам не зависящим от размера базы, но и благодаря небольшим константам в оценках.
- 3)Универсальность, независимость от языка.
- 4)Повышенная точность, благодаря использованию синтаксических связей.

Благодаря данным преимуществам методы без сомнения подходят для решения поставленной задачи.

### Литература

1. Дэн Гасфилд, перевод И.В. Романовского Строки, деревья и последовательности в алгоритмах, Санкт-Петербург 2003
2. Aho,A, Korasick.M. Efficient string matching: an aid to bibliographic search //Comm. ACM. 1975. vol. 18. P. 333-340
3. Sandeep Tata, R. A. Hankins, J. M. Patel Practical Suffix Tree Construction //University of Michigan 1301; Ann Arbor, MI 48109-2122; USA
4. Marie-Catherine de Marneffe, Bill MacCartney and Christopher D. Manning. 2006. Generating Typed Dependency Parses from Phrase Structure Parses. InLREC 2006.
5. Marie-Catherine de Marneffe and Christopher D. Manning. 2008. The Stanford typed dependencies representation. In COLING 2008 Workshop on Cross-framework and Cross-domain Parser Evaluation.

### Слова благодарности

Особую благодарность хочется выразить замечательной девушке Катерине Смышчѐк за неоценимую помощь в лингвистике) Анне Толкачевой, Дмитрию Мокееву и Александру Ванину) Так как мы все вместе думаем над одними и теми же проблемами) И Илье Лысенкову человеку объединившему нас, человеку который всегда поможет умным советом)

## Оценка времени анализа событий системой обнаружения и предотвращения сетевых атак

*Казачкин Дмитрий Сергеевич*

*Аспирант*

*МГУ им. Ломоносова, ВМиК, Москва, Россия*

*email: zok@lvk.cs.msu.su*

Доклад посвящен актуальной в области информационной безопасности задаче анализа сетевого трафика на высокоскоростных сетевых каналах. Эффективность систем защиты от компьютерных атак, поточных антивирусов, межсетевых экранов и подобных систем при функционировании на современных каналах существенно определяется эффективностью решения задач по

сбору данных с канала, передаче их по стеку протоколов в ядре ОС, и собственно задач анализа данных.

Обзор существующих средств анализа сетевого трафика показал, что ни одно из них не имеет встроенных средств оценки эффективности решения данных задач, т.е. оценки наихудшего времени обработки сетевого пакета набором загруженных правил. Обычно системы лишь тестируют в среднем случае (на произвольном трафике), который может существенно отличаться от худшего: в работе [2] приводится пример различия в несколько порядков при анализе одного пакета одним правилом. Чрезмерно долгий анализ приводит к переполнению очереди и, как следствие, к потере части трафика.

Для упрощения построения анализаторов сетевого трафика на факультете Вычислительной математики и кибернетики МГУ разрабатывается специализированная программная среда AURA [1], в состав которой входит проблемно-ориентированный автоматный язык программирования, среда поддержки выполнения программ на этом языке и набор библиотек функций, типичных для реализации алгоритмов анализа.

Основной задачей данного исследования является создание метода приближенной оценки наихудшего времени выполнения заданного набора анализаторов трафика в среде AURA. Сюда входит как время на собственно обработку события правилами анализа, так и на захват пакетов и другие служебные операции.

В докладе описывается формальная модель решаемой задачи, учитывающая специфику автоматного языка среды AURA. Основываясь на формальной модели, предлагается методика получения численных оценок для наихудшего времени выполнения на конкретном оборудовании, сочетающая статический анализ и профилировку. Для функций, типичных для анализа трафика, указываются дополнительные эвристики.

Данный метод позволяет по фиксированному набору правил ответить на вопрос о максимальном уровне загруженности канала, при которой анализатор способен обрабатывать все сетевые пакеты в соответствии с данными правилами без потерь.

## Литература

1. Гамаюнов Д.Ю., Казачкин Д.С., Шугалёв П.И. AURA: среда высокоскоростного анализа сетевого трафика для задач информационной безопасности // Защита информации. Инсайд. 2009. No 4. С. 32-35.
2. Randy Smith, Cristian Estan, Somesh Jha. Backtracking Algorithmic Complexity Attacks Against a NIDS. <http://www.acsac.org/2006/papers/54.pdf>

# Автоматическая генерация тестов по комментариям к программному коду

*Носков Алексей Анатольевич*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,*

*Россия*

*email: alno87@mail.ru*

Для повышения надежности программного обеспечения все большее распространение получают методы, основанные на автоматизированном тестировании кода, но при этом тесты все равно пишутся, как правило, вручную. Обычно вручную составляется также документация к программному коду, являющаяся необходимой компонентой качественного ПО. Однако указанные работы являются достаточно рутинными и часто производятся лишь частично. В то же время комментарии, аналогично тестам, фиксируют некоторые аспекты поведения кода. Кроме того, хорошей практикой является включение в комментарии примеров использования программных единиц – процедур, функций и методов, а также описания вариантов их поведения в зависимости от различных условий. Таким образом, две рутинных работы частично дублируют друг друга.

В представленном докладе предлагается метод автоматической генерации тестов по описанию поведения программных единиц, содержащемуся в комментариях на естественном языке, что позволяет сократить общий объем сопутствующих созданию ПО работ. Для построения тестов метод использует представленную в комментариях информацию о параметрах функций, примерах использования, типах исключений, порождаемых при нарушении предусловий.

Предлагаемый метод основывается на двухэтапном анализе текста комментариев с помощью лексико-синтаксических шаблонов языка LSPL, который используется для формального описания языковых конструкций, с учетом их структуры, лексических и морфосинтаксических характеристик. Для выделения конструкций по шаблонам LSPL были разработаны соответствующие программные средства [1].

На первом этапе анализа с помощью шаблонов именных групп извлекаются возможные именованные элементы кода. Семантика именованных элементов определяется из разметки комментариев для систем автоматической генерации документации (Javadoc, Doxygen, RDoc). Для каждого элемента формируется набор шаблонов для выделения их употреблений в тексте. На втором этапе на основе сформированных шаблонов именованных элементов строятся шаблоны конструкций, описывающих аспекты поведения кода. Эти шаблоны используются для выделения из текста комментариев набора ограничений, описывающих поведение кода, а из них - наборы проверяемых в тестах условий, по которым генерируется код автоматизированных тестов.

Метод был реализован на языке Java в виде модуля для IDE Eclipse и был

опробован для Javadoc-комментариев.

### Литература

1. Носков А.А. Метод выделения в тексте конструкций по их лексико-синтаксическим шаблонам // Сборник статей молодых ученых факультета ВМиК МГУ. М.: Издательский отдел фак-та ВМиК МГУ им. М.В. Ломоносова; МАКС Пресс, 2009. Выпуск 6. С.136-145.

## Интеллектуальная информационная система атрибуции художественных текстов

*Попов Никита Евгеньевич*

*Студент*

*Московский государственный университет экономики, статистики и информатики, Институт Компьютерных Технологий, Химки, Россия*  
*email: nikeropov@rambler.ru*

Стремление расширить область применения цифровых вычислительных машин существует со времени их появления. В известной мере это связано с требованиями практики, с поиском наиболее эффективных способов деятельности. Отчасти это вызвано также вполне понятным стремлением к усовершенствованию конструкции или способов программирования с целью придать машинам новые, ранее недоступные им функции. Так или иначе, обе эти причины имеют отношение к определенному разделу науки об искусственном интеллекте.

Ряд таких задач связан с классификацией объектов — отнесением объектов или явлений к одному из нескольких predetermined классов. Естественно, что для каждой отдельной области применения требуется, в конечном счете, собрать воедино специфические характеристики конкретной задачи.

Основным результатом проведенного исследования является создание интеллектуальной информационной системы для распознавания авторства художественных текстов, написанных на русском или ином языке, использующем фонологическую (неиероглифическую) письменность. При проектировании ИИС были проанализированы существующие методы и информационные системы в выбранной предметной области. Наиболее успешные подходы были использованы при разработке представления базы знаний системы:

- модель последовательности букв текста как реализации цепи А.А. Маркова [1];
- частотный анализ вхождения служебных слов (лингвистический спектр Морозова) [2,3];
- частотный анализ использования знаков пунктуации [4];
- частотный анализ количества букв в слове [4];
- частотный анализ количества слов в предложении [4].

После обучения системы, т.е. ввода в нее отобранных произведений различных жанров (русская классическая литература, современные детективы

и фантастика, поэзия, историческая литература, переводы произведений иностранных авторов) на русском и английском языках (корпус текстов), было проведено перекрестное тестирование.

В процессе тестирования было установлено, что наиболее правильные и стабильные результаты дают метод марковских цепей, основанный на анализе матрицы переходов для пар букв и метод лингвистических спектров Морозова.

### Литература

1. Кукушкина О.В., Поликарпов А.А., Хмелёв Д.В. Определение авторства текста с использованием буквенной и грамматической информации. // Проблемы передачи информации, 2001, т.37, вып.2, с.96-108.
2. Марков А.А. Об одном применении статистического метода // Докладено в заседании Отделения Физико-математических наук 17 февраля 1916 г
3. Морозов Н.А. Лингвистические спектры: средство для отличения плагиатов от истинных произведений того или иного неизвестного автора. Стилеметрический этюд // Известия отд. русского языка и словесности Имп.Акад.наук, т.XX, кн.4, 1915.
4. Хетсо Г. Принадлежность Достоевскому: к вопросу об атрибуции Ф.М. Достоевскому анонимных статей в журналах "Время" и "Эпоха" // SOLUM FORLAG A.S.: OSLO 1986.

## Итерационная модель нейросетевой ассоциативной памяти

*Прокопович Григорий Александрович*

*Аспирант*

*Объединённый институт проблем информатики НАН Беларуси,  
Лаборатория моделирования самоорганизующихся систем, Минск*

*email: bigznich@tut.by*

В [1] предложена простая модель детерминированной хаотической системы, которая способна осуществлять ассоциативную обработку неструктурированной информации, которая записана на предельных циклах одно- или многомерных отображениях специального вида.

Хотя указанная модель эффективно производит обработку информации, включая поиск и восстановление по произвольному фрагменту, она обладает рядом недостатков: 1) при добавлении в память новых образов необходима полная перестройка отображения; 2) информационные блоки кодируются вещественными числами, что влечёт за собой вычислительные затраты.

В связи с этим, для модели [1] предлагается новый оригинальный метод кодирования информационных блоков. Рассмотрим входной вектор  $X$  как

строго определённую последовательность уникальных векторов  $\mathbf{x}_i$  одинакового размера, т.е.  $\mathbf{X}=[\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3]$ . Для примера, на рис. 1 изображены исходные двумерные вектора, содержащие графическую информацию о монохромных изображениях букв "А", "В" и "С", которые разбиты на три равные части таким образом, что каждая из них содержит информацию, присущую только своему исходному вектору.

Предложенный метод заключается в кодировании вектора  $\mathbf{X}$  как замкнутой последовательности пар векторов  $\mathbf{x}_1 \rightarrow \mathbf{x}_2$ ,  $\mathbf{x}_2 \rightarrow \mathbf{x}_3$  и  $\mathbf{x}_3 \rightarrow \mathbf{x}_1$  с помощью нейросетевой гетероассоциативной памяти ДАП [2]. В результате, на основе сети ДАП получается отображение специального вида:  $\mathbf{x}_{i+1} = F(\mathbf{x}_i)$ , причём,  $\mathbf{x}_n = F(\mathbf{x}_1)$ , где  $n$  — количество информационных блоков, на которые был разбит первоначальный вектор  $\mathbf{X}$ , в нашем случае  $n=3$ . Таким образом, носителями информации являются предельные циклы фазового пространства состояний сети ДАП.

После проведённого процесса обучения, сеть ДАП способна восстанавливать и идентифицировать вектор  $\mathbf{X}$  по любому известному ключевому вектору  $\mathbf{x}_i$ . Так как итерационный процесс извлечения информации является циклическим, то предложенный метод регламентирует и правило останова. Процесс распознавания считается успешно завершённым, если за конечное число итераций на выходе сети ДАП появится вектор, идентичный ключевому. Если описанное правило не выполняется, то считается, что начальная точка не попала на аттрактор в фазовом пространстве состояний сети ДАП.

Предложенная итерационная модель нейросетевой ассоциативной памяти позволяет эффективно оперировать неструктурированной информацией. Благодаря своей нейросетевой архитектуре разработанная модель памяти, по сравнению с [1], имеет ряд преимуществ: простой способ кодирования информационных блоков и вычисления векторного отображения, а также более быстрый способ добавления и удаления образов, не требующий перебор и сортировку блоков информации.

## Литература

1. Андреев Ю.В., Дмитриев А.С., Куминов Д.А. Хаотические процессоры // Успехи современной радиоэлектроники. 1997. №10. С. 50-79.
2. Прокопович Г.А. Адаптивный нейросетевой классификатор // Информатика. 2009. №3(23). С. 68-81.

## Иллюстрации





Рис. 1. Предварительная подготовка векторов к записи.

## Способ моделирования XML-данных выражениями языка C++

*Струков Илья Викторович*  
*Студент (специалист)*

*Московский государственный университет имени М.В. Ломоносова*  
*Факультет вычислительной математики и кибернетики, Москва,*  
*Россия*

*email: iley@iley.ru*

Сегодня широко распространён XML [1] — расширяемый язык разметки документов.

Метод непосредственной интеграции [3] заключается в том, чтобы средствами основного языка программирования задавать выражения, семантически эквивалентные и синтаксически близкие конструкциям альтернативного языка. На таком принципе основана, например, библиотека *InteLib*. Она позволяет встраивать код на языке *Lisp* в программы на *C++*. Данный доклад посвящён применению аналогичного подхода к моделированию XML-данных.

Рассмотрим пример. Слева представлены некоторые XML-данные в традиционном представлении, справа — соответствующая им запись в виде выражений на языке *C++*.

	<code>data = xml</code>
<code>&lt;html&gt;</code>	<code>(html)</code>
<code>  &lt;body&gt;</code>	<code>(body)</code>
<code>  &lt;h1 align="center"&gt;</code>	<code>(h1, align="center")</code>
<code>    Title</code>	<code>  + "Title"</code>
<code>&lt;/h1&gt;</code>	<code>(!h1)</code>
<code>&lt;p&gt;</code>	<code>(p)</code>
<code>  Some text here...</code>	<code>  + "Some text here..."</code>
<code>&lt;/p&gt;</code>	<code>(!p)</code>
<code>&lt;/body&gt;</code>	<code>(!body)</code>
<code>&lt;/html&gt;</code>	<code>(!html);</code>

Чтобы использовать приведённый код в реальной программе, необходимо

заранее определить объекты `data`, `xml`, `html` и т.д. Объекты `html`, `body`, `h1` и `p` представляют имена XML-элементов. Они должны обеспечивать перегрузку оператора `!`. `align` представляет имя атрибута, для него должен быть перегружен оператор `=`. Объект `xml` служит для того, чтобы объединить теги и построить объект, представляющий всю XML-структуру. Для него должны быть перегружены операторы `()` и `+`. О возможности перегрузки операторов в C++ см. [2].

### Литература

1. Гарольд Э., Минс С. XML. Справочник. СПб.: Символ-Плюс. 2002.
2. Страуструп Б. Язык программирования C++. М.: Бином. 2004.
3. E. Bolshakova, A. Stolyarov. Building functional techniques into an object-oriented system. // Proceedings of the 4th JCKBSE, vol.62 of Frontiers in Artificial Intelligence and Applications. Amsterdam. IOS Press. 2000. С. 101-106.

### Слова благодарности

Автор выражает признательность своему научному руководителю Столярову Андрею Викторовичу.

## Система создания и управления мультимедийными документами Fuzzle CMS

*Январев Владислав Игоревич*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Королев,  
Россия*

*email: janvarev@gmail.com*

Мультимедийный документ, характерным примером которого является современный интернет-сайт, обычно включает разнородные объекты: текст, изображение, звук и видео. Фрагменты такого документа могут обладать интерактивным поведением.

Основная сложность создания мультимедийного интернет-сайта связана с интеграцией разнородных объектов и задания их интерактивного поведения, что требует участия разных специалистов - верстальщиков, программистов, администраторов.

В настоящее время существует несколько классов программных инструментов для создания мультимедийных документов: CMS для HTML-сайтов (Bitrix и пр.); CMS для Flash-сайтов, или Flash CMS (FlashMoto и пр.); настольные редакторы, к которым можно отнести Adobe Flash или Microsoft PowerPoint. В системах всех этих классов возможности интеграции разнородных интерактивных объектов либо ограничены, либо слишком сложны для неподготовленного пользователя.

В докладе рассмотрена разработанная система управления Flash-сайтами Fuzzle CMS, предлагающая более гибкие возможности интеграции для подготовки современного мультимедийного интернет-сайта. При разработке системы были использованы следующие принципы:

Модульность – возможность динамического программного подключения к документу новых видов интерактивных объектов;

Проектирование пользовательского интерфейса с учетом расширяемости и доступности для неподготовленного пользователя;

Возможность коллективной работы – многопользовательский режим, сохранение и сравнение версий документа, просмотр изменений версий в визуальном режиме.

Исходя из этих принципов предложено внутреннее представление мультимедийного документа, в котором каждый визуальный объект документа состоит из трех частей: данных, предназначенных для визуализации, и двух динамически подключаемых программных компонентов, обеспечивающих соответственно визуализацию и редактирование этих данных. Благодаря такому внутреннему представлению разные объекты документа могут представляться и управляться одинаковым образом. Новые виды объектов добавляются в систему путем регистрации в системе их программных компонентов и сразу становятся доступны пользователю.

Единообразное внутреннее представление в мультимедийном документе объектов разных видов позволило реализовать сравнение его версий и визуализировать результаты этого сравнения.

Проведенное сопоставление возможностей системы Fuzzle CMS с другими классами систем управления мультимедийными документами показало, что она не уступает им в отношении простоты использования (доступна неподготовленному пользователю), и в то же время превосходит их в организации механизмов расширяемости и коллективной работы. С момента создания системы Fuzzle CMS в марте 2009 г. было создано около 20 интернет-сайтов, часть из которых разработана пользователями-неспециалистами, не владеющими навыками программирования и создания сайтов.

Примеры реализованных интернет-сайтов, демонстрационная версия системы и документация к ней доступны на сайте <http://fuzzle-cms.ru/>.

# Теория вероятностей и математическая статистика

## Асимптотические оценки абсолютной постоянной в неравенстве Берри-Эссеена для распределений, не имеющих третьего момента

*Гапонова Маргарита Олеговна*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова  
Факультет вычислительной математики и кибернетики, Москва,*

*Россия*

*email: mail@gmail.com*

Пусть  $X_1, \dots, X_m, \dots$  — последовательность независимых одинаково распределенных случайных величин, удовлетворяющих соотношениям

$$EX_1 = 0, \quad DX_1 = 1, \quad E|X_1|^{2+\delta} \equiv \beta_{2+\delta} < \infty,$$

для некоторого  $0 < \delta \leq 1$ . Обозначим

$$S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i, \quad F_n(x) = \mathbf{P}(S_n < x), \quad \Phi(x) = \int_{-\infty}^x e^{-t^2/2} dt$$

В этом случае справедливо неравенство Берри-Эссеена (см. [4,5] для  $\delta = 1$  и [3] для  $0 < \delta \leq 1$ )

$$\rho(F_n, \Phi) \equiv \sup_x |F_n(x) - \Phi(x)| \leq C(\delta) \cdot \varepsilon,$$

где  $\varepsilon = \frac{\beta_{2+\delta}}{n^{\delta/2}}$  — дробь Ляпунова,  $C(\delta)$  зависит только от  $\delta$ .

Известно, что  $C(1) \leq 0.4784$  [2]. Наилучшие на сегодняшний день верхние оценки  $C(\delta)$  для  $0 < \delta < 1$  получены в работе [1]. На практике  $\varepsilon$  мало, поэтому интересен вопрос об асимптотических оценках  $C(\delta)$ , которые впервые найдены Х. Правитцем [6] для  $\delta = 1$  :

$$C(1) \leq \begin{cases} 0.5151, & \varepsilon \leq 0.0985, \\ 0.5270, & \varepsilon \leq 0.1387. \end{cases}$$

В данной работе обобщен результат Правитца на случай  $0 < \delta < 1$ . Получена оценка вида

$$\rho(F_n, \Phi) \leq C_1(\ell, \delta) \cdot \varepsilon + C_2(\ell, \delta) \cdot n^{-\delta/2} \text{ при } \frac{\beta_{2+\delta} + 1}{n^{\delta/2}} \leq 2\ell.$$

Значения величин  $C_1(\ell, \delta)$ ,  $C_2(\ell, \delta)$  и  $C(\delta)$  для некоторых  $\ell$  и  $\delta$  представлены в следующей таблице:

$\delta$	$\ell$					
	0.3			0.05		
	$C_1$	$C_2$	$C$	$C_1$	$C_2$	$C$
0.1	0.1500	0.0021	0.1521	0.1480	0.0001	0.1481
0.2	0.2915	0.0327	0.3242	0.1276	0.0001	0.1277
0.3	0.2776	0.0714	0.3490	0.1122	0.0010	0.1132
0.4	0.2765	0.1044	0.3809	0.1044	0.0064	0.1108
0.5	0.2800	0.1382	0.4182	0.1075	0.0200	0.1275
0.6	0.2852	0.1655	0.4507	0.1221	0.0431	0.1652
0.7	0.2912	0.1921	0.4833	0.1471	0.0747	0.2218
0.8	0.2987	0.2166	0.5053	0.1853	0.1132	0.2985
0.9	0.3085	0.2399	0.5484	0.2250	0.1568	0.3818
1.0	0.3254	0.2624	0.5878	0.2727	0.2041	0.4768

### Литература

1. М. Е. Григорьева, И. Г. Шевцова. Уточнение неравенства Каца–Берри–Эссеена.// Информатика и ее применения. 2010. Т.4. В печати.
2. В. Ю. Королев, И. Г. Шевцова. Уточнение неравенства Берри–Эссеена с приложениями к пуассоновским и смешанным пуассоновским случайным суммам.// Обозрение прикладной и промышленной математики, 2010, в печати.
3. В. В. Петров. Суммы независимых случайных величин. Москва, Наука, 1972.
4. А. С. Berry. The accuracy of the Gaussian approximation to the sum of independent variates.// Trans. Amer. Math. Soc., 1941, vol.49, p.122–139.
5. С. G. Esseen. On the Liapunoff limit of error in the theory of probability.// Ark. Mat. Astron. Fys., 1942, vol.A28, No.9, p.1-19.
6. Н. Prawitz. On the remainder in the central limit theorem.// Scand. Actuarial J., 1975, No.3, pp.145-156.

# Уточнение неравенства Каца–Берри–Эссеена

*Григорьева Мария Евгеньевна*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова*

*Факультет вычислительной математики и кибернетики*

*Москва, Россия*

*email: maria-grigoryeva@yandex.ru*

Для  $0 \leq \delta \leq 1$  обозначим через  $\mathcal{F}_{2+\delta}$  множество функций распределения с нулевым средним, единичной дисперсией и конечным абсолютным моментом  $\beta_{2+\delta}$  порядка  $2 + \delta$ . Пусть  $X_1, \dots, X_n$  – независимые одинаково распределенные случайные величины с общей функцией распределения  $F \in \mathcal{F}_{2+\delta}$ ,  $F_n(x) = F^{*n}(x\sqrt{n}) = \mathbb{P}\left(\frac{X_1 + \dots + X_n}{\sqrt{n}} < x\right)$ ,  $\Phi(x)$  – функция стандартного нормального распределения. В сделанных выше предположениях оценка скорости сходимости в центральной предельной теореме устанавливается неравенством Каца–Берри–Эссеена [2] согласно которому

$$\rho(F_n, \Phi) \equiv \sup_x |F_n(x) - \Phi(x)| \leq C_0(\delta) \cdot \frac{\beta_{2+\delta}}{n^{\delta/2}}, \quad (1)$$

для всех  $n \geq 1$  и  $F \in \mathcal{F}_{2+\delta}$ .

Верхние оценки величины  $C_0(\delta)$  при некоторых  $0 < \delta < 1$  впервые были построены в 1983 году В. Тысяком [3], и недавно были уточнены М. О. Гапоновой, А. Ю. Корчагиным и И. Г. Шевцовой [1].

В данной работе уточняются верхняя оценка константы  $C_0(\delta)$  в (1), а также верхняя оценка константы  $C_1(\delta)$  в неравенстве

$$\rho(F_n, \Phi) \leq C_1(\delta) \cdot \frac{\beta_{2+\delta} + 1}{n^{\delta/2}}, \quad n \geq 1, F \in \mathcal{F}_{2+\delta}. \quad (2)$$

Решение последней задачи играет принципиальную роль в уточнении констант в аналоге неравенства Каца–Берри–Эссеена для пуассоновских и смешанных пуассоновских случайных сумм.

Полученные оценки приведены в следующей таблице:

$\delta$	$C_0$	$C_1$	$\delta$	$C_0$	$C_1$
0.9	0.5383	0.3088	0.4	0.6342	0.4079
0.8	0.5723	0.3187	0.3	0.6195	0.4450
0.7	0.6026	0.3334	0.2	0.6094	0.4901
0.6	0.6276	0.3528	0.1	0.6028	0.5451
0.5	0.6413	0.3775			

## Литература

1. М. О. Гапонова, А. Ю. Корчагин, И. Г. Шевцова. Об абсолютных константах в равномерной оценке точности нормальной аппроксимации для распределений, не имеющих третьего момента // Сборник статей молодых ученых ВМК МГУ, 2009, вып. 6, с. 81–89.

2. В. В. Петров. Суммы независимых случайных величин. Москва, “Наука”, 1972.
3. W. Tysiak. Gleichmäßige und nicht-gleichmäßige Berry-Esseen-Abschätzungen//  
Dissertation, Wuppertal, 1983.

### Слова благодарности

Автор выражает признательность И.Г. Шевцовой и В.Ю.Королеву за неизменное внимание и поддержку.

## Исследование применимости моделей оценки надёжности для разработки программного обеспечения с открытым исходным кодом.

*Зорин Даниил Александрович*

*Студент*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: jfaranime@gmail.com*

Моделями оценки надежности (англ. Software reliability growth models) называются специальные математические модели, позволяющие по статистике тестирования за определенный период времени оценить количественные характеристики надежности проекта. Несмотря на то, что в литературе рассмотрено множество моделей оценки надёжности, критерии выбора этих моделей для конкретного проекта не определены четко. Соответственно, для применения моделей оценки надежности и получения адекватных результатов необходимо найти способ сравнивать модели между собой и выбирать наиболее подходящую.

Выбор моделей невозможен в отрыве от конкретной предметной области. В настоящее время широко распространена разработка программного обеспечения с открытым исходным кодом. Автору известно лишь две работы, посвященных проблеме выбора модели оценки надежности для таких проектов.

В данной работе рассматриваются различные модели оценки надежности, с целью определения применимости моделей к проектам с открытым исходным кодом и уточнения механизма их использования описываются эксперименты с некоторыми крупными проектами с открытым исходным кодом. В результате экспериментов получены различные характеристики надежности рассматриваемых проектов и сформулированы гипотезы об общих свойствах проектов с открытым исходным кодом.

### Литература

1. M.Lyu Software Reliability Engineering: A Roadmap // 2007 Future of Software Engineering. Washington, DC, USA, 2007. p. 153 — 170.
2. Ширяев А. Н. Вероятность. М.: МЦНМО, 2004.
3. R.Al-Ekram Software Reliability Growth Modeling and Prediction. Waterloo: Dept. of Electrical and Computer Engineering, University of Waterloo, 2002
4. S.Alam Software Reliability Using Markov Chain Usage Model. Dhaka, Bangladesh: Department of Computer Science and engineering Bangladesh University of Engineering and Technology, 2005
5. A.Wood Software Reliability Growth Models. Technical Report, 1996
6. H.Umeda The development of a program drawing reliability growth curves and implementing reliability growth models. Tottori University of Natural Sciences, Department of Information Systems, 2008
7. M.Limnios, N.Nikulin Recent Advances in Reliability Theory. Birkhauser, 2000
8. C.Stringfellow An Empirical Method for Selecting Software Reliability Growth Models // Empirical Software Engineering 2002. 7. №4. p. 319 — 343.

## **Классификация многоканальных дистанционных изображений с использованием марковских**

**случайных полей и копул  
Крылов Владимир Андреевич**

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,*

*Россия*

*email: vl\_krylov@mail.ru*

Одним из важнейших источников информации при решении задач современной картографии, риск менеджмента (пожары, наводнения), эпидемиологии являются спутники и радары. Современные орбитальные системы предоставляет возможность получать многоканальные изображения высокого разрешения, число каналов в которых в некоторых случаях доходит до нескольких десятков. Для обработки таких данных требуется адаптация существующих методов одноканальной обработки для работы с многоканальными изображениями. В данной работе рассматривается один из возможных статистических подходов к решению одной из базовых задач обработки дистанционных изображений - задаче классификации, в случае многоканальных изображений.



Базовая идея предлагаемого метода состоит в моделировании плотностей распределений отдельных каналов при помощи традиционных методов [3], с последующим построением совместного распределения с использованием копул [4]. Для борьбы с шумом, неизбежно присутствующим на дистанционных изображениях, предлагается использовать контекстную модель марковских случайных полей [1]. Полученная в результате модель имеет распределение Гиббса [2] с некоторой энергией, которая оптимизируется с использованием статистического метода имитации отжига [2] на заданном изображении. Рассмотренные эксперименты с изображениями, полученными радаром с синтезированной апертурой, позволяют констатировать превосходство предложенной модели в сравнении с классическими методами (использование многомерных Wishart и Накагами распределений [3]).

### Литература

1. Besag J. On the statistical analysis of dirty pictures // Journal of the Royal Statistical Society B, Vol. 48. P. 259-302, 1986.
2. Geman S., Geman D. Stochastic relaxation, Gibbs distributions and the Bayesian restoration // IEEE Trans. Pattern Anal. Mach. Intell., vol. 6, no. 6, pp. 721-741, 1984.
3. Landgrebe D. A. Signal Theory Methods in Multispectral Remote Sensing. New-York: Wiley, 2003.
4. Nelsen R. B. An Introduction to Copulas. New-York: Springer, 2007.

## Асимптотическая нормальность оценок риска пороговой обработки вейвлет-коэффициентов

*Маркин Артём Васильевич*

*Соискатель*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: artem.v.markin@mail.ru*

Рассматривается пороговая обработка вейвлет-коэффициентов при прямом наблюдении одномерного сигнала и пороговая обработка вейвлет-вейвлет коэффициентов в задаче томографии при наблюдении двумерного сигнала через оператор Радона.

В одномерном случае наблюдается сигнал, состоящий из полезного сигнала и гауссовского шума. К сигналу применяется дискретное вейвлет-преобразование (Mallat, 1999), а затем к коэффициентам деталей применяется мягкая или жесткая пороговая обработка с универсальным порогом (Donoho, 1994). Для неизвестного риска такой операции строится оценка риска. В работе показывается, что при определенных ограничениях на регулярность сигнала и на оценку дисперсии шума указанная оценка риска обладает свойством асимптотической нормальности.

В двумерном случае в задаче томографии наблюдается не сам сигнал, а его радоновских образ (Наттерер, 1990) с аддитивным гауссовским шумом. Задача томографии, т. е. задача обращения оператора Радона, решается с помощью вейвлетов и функций специального вида, называемых вейглетами (Донoho, 1995). К радоновскому образу сигнала применяется вейвлет-преобразование, полученные коэффициенты являются вейвлет-коэффициентами искомого полезного сигнала. В силу некорректности задачи томографии дисперсия вейвлет-вейвлет коэффициентов растет при измельчении масштаба разложения. К коэффициентам применяется мягкая пороговая обработка с порогом Колашика (Kolaczyk, 1996), которая выполняет роль регуляризации. Оценка риска пороговой обработки также будет обладать свойством асимптотической нормальности. Проводится сравнение с одномерным случаем и анализируется влияние некорректности задачи томографии.

### Литература

1. Наттерер Ф. Математические аспекты компьютерной томографии. М., 1990.
2. Donoho D. L., Johnstone I. M. Ideal spatial adaptation via wavelet shrinkage // *Biometrika*, 1994. Vol. 81. No. 3. P. 425–455.
3. Donoho D. L. Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition // *Applied and Computational Harmonic Analysis*, 1995. Vol. 2. P. 101–126.
4. Kolaczyk E. D. A wavelet shrinkage approach to tomographic image reconstruction // *J. Amer. Statist. Assoc.*, 1996. Vol. 91. No. 435. P. 1079–1090.
5. Mallat S. A wavelet tour of signal processing. Academic Press, 1999.

## Численный поиск нижней оценки для абсолютной константы в неравенстве Берри-Эссеена для пуассоновских случайных сумм.

*Нефедова Юлия Сергеевна*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: julia\_n@inbox.ru*

Пусть  $X_1, X_2, \dots$  – последовательность независимых одинаково распределенных случайных величин с функцией распределения  $F(x) = \mathbf{P}(X_1 < x)$  и таких, что:

$$\mathbf{E}X_1 = 0, \quad \mathbf{D}X_1 = 1 \quad \beta_3 \equiv \mathbf{E}|X_1|^3 < \infty. \quad (1)$$

Пусть  $N_\lambda$  – случайная величина, имеющая распределение Пуассона с параметром  $\lambda > 0$ . Предположим, что при каждом  $\lambda > 0$  случайные величины  $N_\lambda, X_1, X_2, \dots$  независимы.

Рассмотрим пуассоновскую случайную сумму  $S_{N_\lambda} = X_1 + \dots + X_{N_\lambda}$ . Для определенности полагаем  $S_{N_\lambda} = 0$  при  $N_\lambda = 0$ .

Известно, что при приведенных выше условиях (1) на моменты случайной величины  $X_1$  справедливо неравенство Берри–Эссеена для пуассоновских случайных сумм: существует абсолютная положительная и конечная постоянная  $C$  такая, что

$$\Delta \equiv \sup_x \left| \mathbf{P} \left( \frac{S_\lambda}{\sqrt{\lambda}} < x \right) - \Phi(x) \right| \leq \frac{C\beta_3}{\sqrt{\lambda}}. \quad (2)$$

Впервые оценка (2) была доказана в 1972 году и опубликована с константой  $C = 2,23$ . В [2] приводится доказательство этого неравенства с наилучшей на сегодняшний день константой  $C = 0,3041$ . Однако нижние оценки для  $C$  найдены не были.

Данная работа имеет своей целью восполнить этот пробел.

**ТЕОРЕМА.** *Для константы  $C$  в неравенстве (2) справедлива следующая оценка*

$$C \geq 0,2344.$$

Рассматривался класс  $\mathcal{F}_2$  – класс функций распределения случайных величин, принимающих два значения и удовлетворяющих моментным условиям (1).

Обозначим

$$C(\lambda) = \sup_{\mathcal{F}_2} \frac{\Delta\sqrt{\lambda}}{\beta_3}.$$

Составлена программа вычисления  $C(\lambda)$ . Результатом работы программы является приведенный ниже график зависимости  $C(\lambda)$  от  $\lambda$ .

Как видно из графика, для константы  $C$  из неравенства (2) справедлива следующая нижняя оценка  $C \geq C_1 = \max_{\lambda > 0} C(\lambda) = C(0,79) = 0,2344$ .

### Литература

1. Золотарев В.М. Современная теория суммирования независимых случайных величин. М.: Наука, 1986.
2. Королев В.Ю., Шевцова И.Г. Уточнение неравенства Берри–Эссеена с приложениями к пуассоновским и смешанным пуассоновским случайным суммам // Обозрение прикладной и промышленной математики, М, 2010.

### Иллюстрации

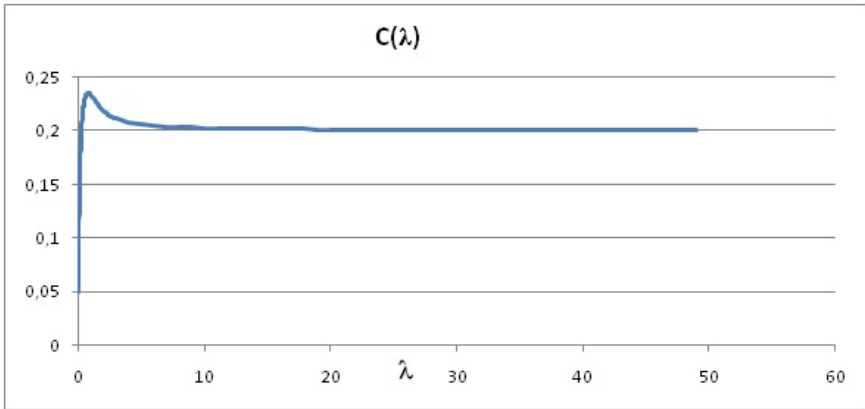


Рис. 1. График зависимости  $C(\lambda)$  от  $\lambda$ .

## Оптимальный выбор числа допустимых ошибок первого рода при множественной проверке гипотез

*Рябенко Евгений Алексеевич*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: riabenko.e@gmail.com*

Проблема множественной проверки гипотез, возникающая в большом количестве исследований в самых различных областях, является одной из наиболее актуальных задач статистического анализа данных. В случаях, когда эксперимент требует покомпонентных выводов по большому числу переменных, игнорирование эффекта множественных сравнений может привести к неконтролируемому росту числа ошибок первого рода и значительному искажению результатов.

Рассмотрим выборку  $X$  из многомерного распределения  $P \in \Omega$  (возможно, неизвестного, или известного с точностью до параметров). Пусть для проверки каждой покомпонентной гипотезы  $H_i : P \in \omega_i, i = 1, \dots, s$ , против общей альтернативы  $H'_i : P \notin \omega_i$  имеется статистика  $T_i$ . Обозначим множество индексов верных нулевых гипотез за  $I(P)$ . Одна из мер, распространяющих на многомерный случай понятие ошибки первого рода, — обобщённая групповая вероятность ошибки ( $k$ -FWER), определяемая как вероятность  $k$  и более ложных отклонений нулевых гипотез:

$$k\text{-FWER} = P\{\text{отвергнуто не менее } k \text{ гипотез } H_i : i \in I(P)\}.$$

Контроль обобщённой групповой вероятности ошибки на уровне  $\alpha$  означает  $k\text{-FWER} \leq \alpha$  для всех  $P$ . Допущение до  $k - 1$  ошибок первого рода позво-

ляет увеличить мощность статистических процедур, обеспечивающих такой контроль. Для малых  $k$  с ростом числа проверяемых гипотез процедуры становятся слишком консервативными, не позволяя обнаружить большую часть отклонений от  $H_0$ .

В статистической литературе величина  $k$  рассматривается как фиксированная до начала анализа, однако, варьируя  $k$ , можно добиться более чётких выводов. Предлагается использовать значение

$$k_{opt} = \min_k \arg \max_k f(P, k) = \min_k \arg \max_k (|R_k(P, \alpha)| - k + 1),$$

где для заданных  $k$  и  $\alpha$   $R_k(P, \alpha)$  — множество индексов отвергаемых гипотез. Вычисляя значения  $k_{opt}$  для различных процедур и одного  $\alpha$ , можно непосредственно сравнивать их консервативность, что может быть достаточно нетривиальной задачей при сопоставлении матриц модифицированных  $p$ -величин, доминирование между которыми может отсутствовать. Кроме того, если статистическая процедура, контролирующая  $k$ -FWER, корректна, то значение  $f_L = f(P, k_{opt})$  можно рассматривать как нижнюю оценку неизвестного числа ложных нулевых гипотез.

Применение описанной идеи продемонстрировано как на модельных данных, так и на примере реальной задачи из области генетики.

### Литература

1. Romano, J.P., Wolf, M. Control of Generalized Error Rates in Multiple Testing // The Annals of Statistics. 2007. No. 35(4). P. 1378-1408.
2. Ge, Y., Dudoit, S., Speed, T.P. Resampling-based multiple testing for microarray data analysis // TEST. 2003. No. 12(1).
3. Golub, T.R., et al. Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring // Science. 1999. No. 286.

## Вейвлет-обработка цветных теневых картин аэродинамического эксперимента

*Хазиахметов Максим Шамилович*

*Студент (специалист)*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,*

*Россия*

*email:*

Понятие "вейвлет" появилось в математике сравнительно недавно. На данный момент разработана мощная математическая теория для их использования в решении различных задач, например, задач обработки сигналов.

В отличие от анализа Фурье теория вейвлетов предлагает технику обработки, при которой возможна четкая локализация изменений сигнала, в то время как коэффициенты Фурье отражают структуру сигнала на всей его

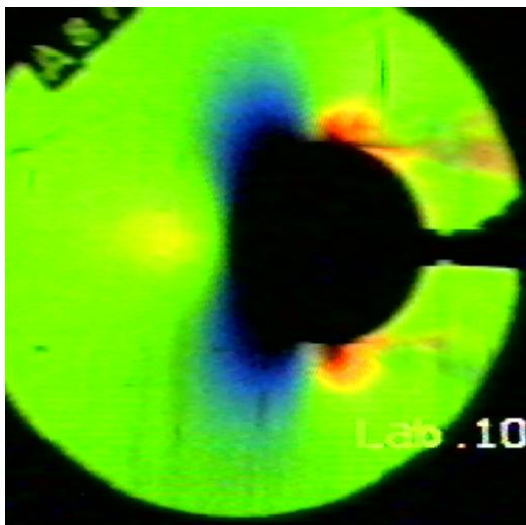


Рис. 1. Исходное изображение.

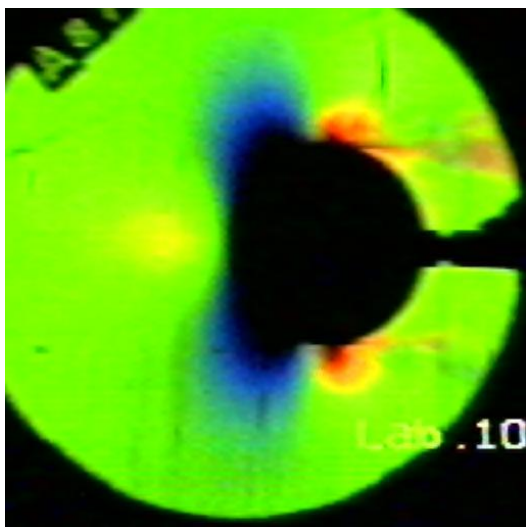


Рис. 2. Изображение после обработки алгоритмом (Вейвлет - gbio1.3, глубина разложения - 2, мягкая обработка универсальным порогом).

протяженности. Вейвлет-преобразование сводится к применению фильтров с необходимыми свойствами: сигнал разделяется на аппроксимирующую и детализирующую составляющие. Различные способы обработки полученного разложения позволяют добиться заданных свойств в восстановленном по измененному набору коэффициентов сигнале.

В настоящее время исследование в области аэродинамики бывают связаны с изучением цветных картин обтекания различных тел: элементов фюзеляжа, сегментов профиля крыла. Изображения, полученные в ходе экспериментов, требуют дополнительной обработки по удалению шумовых эффектов и вычислению истинных цветов. Задача надлежащей математической обработки здесь стоит весьма остро, т. к. особый интерес в настоящее время представляют малые флуктуации характеристик потока. Один из подходов к решению указанных проблем был реализован т. н. дискриминантным методом. Возник вопрос о решении той же задачи методами вейвлет-анализа.

Предложенное решение заключается в том, что для детализирующих коэффициентов применяется т. н. пороговая обработка. Ее описывают следующие формулы преобразования коэффициентов:

$$F(x) = \begin{cases} x, & \text{если } |x| > T, \\ 0, & \text{если } |x| \leq T \end{cases} \quad - \text{ для жесткой обработки;}$$

$$F(x) = \begin{cases} x - T, & \text{если } x > T, \\ x + T, & \text{если } x < -T, \\ 0, & \text{если } |x| \leq T \end{cases} \quad - \text{ для мягкой.}$$

$T$  - порог, он выбирается исходя из принципа минимизации риска оценки истинного значения. Существует много методов его выбора, использующих статистические параметры сигнала.

В настоящий момент реализован алгоритм, который действует следующим образом:

1. Вейвлет-разложение изображения по любой комбинации входных параметров {базис разложения, глубина разложения};
2. Вычисление порогов для каждого уровня разложения, направления деталей и цветовой компоненты;
3. Пороговая обработка коэффициентов (мягкая или жесткая);
4. Восстановление изображения на основе обработанных коэффициентов.

Изображения, полученные с его помощью, визуально мало отличаются от оригинальных, вместе с тем прослеживается сглаживание, характер которого зависит от параметров разложения и типа обработки, что свидетельствует об удалении шумов и возможности настройки алгоритма. Энергии исходного и результирующего изображений, отличаются на сотые доли процента (зависит от настроек), что указывает на необходимость точного подбора параметров разложения.

В дальнейшем предполагается реализовать дополнительные методы выбора порога, обработку т. н. пакетным вейвлет-разложением, выделение областей с заданными свойствами и их границ.

## Метод разумных целей для многокритериальных задач выбора с вероятностными данными

*Холмов Алексей Владимирович*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,*

*Россия*

*email: axeluy@yandex.ru*

В методе разумных целей (МРЦ), предложенном в [1] для решения конечных многокритериальных задач выбора, альтернативам  $X = (x^1, \dots, x^N)$  соответствуют точки  $Y = (y^1, \dots, y^N)$  пространства  $\mathbb{R}^m$ . При  $y \rightarrow \max$  аппроксимируется оболочка Эджворта-Парето (ОЭП) выпуклой оболочки  $Y : Y_P^C = \{y \in \mathbb{R}^m : y \leq y', y' \in \text{conv}Y\}$ . Для визуализации паретовой границы в [2] был разработан метод Диалоговых карт решений (ДКР), с помощью которого пользователь может указать целевую точку  $y^*$ , а затем найти несколько "близких" к выбранной цели альтернатив. Если значения критериев альтернативы  $x^k$  задаются вероятностными распределениями  $(\xi_1^k, \dots, \xi_m^k)$ , имеющие ограниченный носитель, которые при этом являются независимыми попарно и в совокупности, то действия аналогичные, но аппроксимируется ОЭП множества  $\text{conv}\bar{Y} = \text{conv} \bigcup_{k=1}^N \bar{y}^k$ , где  $\bar{y}^k = (\text{ess sup } \xi_1^k, \dots, \text{ess sup } \xi_m^k)$ .

Пусть теперь  $Y = \bigcup_{k=1}^N \text{supp } \xi^k, \mathcal{K}$  — непустой подкласс класса строго монотонных функций на  $Y$  с максимумом в точке  $y^*$ . Зафиксируем некоторое число  $\alpha \in [0, \frac{1}{2})$ . Обобщенно близкими вариантами из  $X$  к цели  $y^*$  относительно класса функций  $\mathcal{K}$  назовем множество  $C(X, \mathcal{K}) = \{x' \in X : \exists U'(\cdot) \in \mathcal{K}, \text{ где } \mathbb{P}(U'(\xi') \geq U'(\xi)) > \alpha, \forall x \in X\}$  (здесь и далее альтернативе  $x$  соответствует случайный вектор  $\xi$ , альтернативе  $x'$  — вектор  $\xi'$ ).

Так как количество  $U'(\cdot)$  в  $\mathcal{K}$  может быть бесконечным, то для построения этого множества "в лоб" приходилось бы решать бесконечное число задач. Бинарное отношение  $\mathcal{R}$  назовем соответствующим  $(X, \mathcal{K})$ , если  $x' \mathcal{R} x \Leftrightarrow \mathbb{P}(U(\xi') > U(\xi)) \geq 1 - \alpha, \forall U(\cdot) \in \mathcal{K}$ .

**Теорема** Если  $\mathcal{R}$  соответствует  $(X, \mathcal{K})$ , то  $C(X, \mathcal{K}) \subseteq \mathbb{K}_{\mathcal{R}}(X)$ , где  $\mathbb{K}_{\mathcal{R}}(X)$  — множество  $\mathcal{R}$ -максимальных элементов на множестве  $X$ .

### Литература

1. Гусев Д.В., Лотов А.В. Методы поддержки принятия решений в задаче конечного выбора // Исследование операций (под ред. Ю.П.Иванилова). М: ВЦ РАН, 1994, с. 15-43.



2. Лотов А.В., Бушенков В.А., Каменев Г.К. и Черных О.Л. Компьютер и поиск компромисса. Метод достижимых целей. М: Наука, 1997.

### Слова благодарности

Работа выполнена при частичной финансовой поддержке гранта Президента РФ по государственной поддержке ведущих научных школ (проект №НШ-2982.2008.1), РФФИ (проект №07-01-00472), программы фундаментальных исследований РАН №14 и программы фундаментальных исследований ОМН РАН №3.

## Распределение вероятностей локального наклона фазового рельефа для задачи развертки фазы в космической радиолокационной топографической интерферометрии

*Шувалов Роман Игоревич*

*Аспирант*

*Московский государственный технический университет имени Н.Э.*

*Баумана, Аэрокосмический, Москва, Россия*

*email: Shuvalov.R.BMSTU@mail.ru*

Большое прикладное значение имеет метод построения цифровых моделей рельефа (ЦМР) поверхности Земли по данным интерферометрической съемки из космоса радиолокатором с синтезированной апертурой антенны (РСА). Это связано с возможностью получать радиолокационные изображения независимо от времени суток и погодных условий; с оперативностью получения ЦМР на интересующий район; с высокой потенциальной точностью метода. Наиболее сложным этапом интерферометрической технологии является развертка фазы [3], так как на получаемых интерферограммах практически всегда присутствуют разрывы. Для каждого разрыва фазы имеется неопределенность его величины и положения. Общий метод выяснения положения системы разрывов заключается в построении распределения вероятностей на множестве всех допустимых систем и поиске его максимума [1].

В настоящей работе построено распределение вероятностей локального наклона фазового рельефа на топографической интерферограмме, получаемой методом космической радиолокационной интерферометрии. Построенное распределение, по сравнению с известными ранее результатами [1, 2, 4], наиболее полно учитывает имеющуюся информацию. Учтены характеристики фазового шума, наблюдаемый наклон фазового рельефа, интенсивность принятого радиолокационного сигнала, априорное распределение вероятностей топографического градиента, возможность наличия на интерферограмме как разрывов фазы обусловленных фазовым шумом, так и разрывов обусловленных рельефом подстилающей поверхности. Разработанное распределение позволяет вычислять вероятности разрывов фазы любой кратности в каждой точке интерферограммы и необходимо для повышения точности построения цифровых моделей рельефа интерферометрическим методом. Проведен вычисли-

тельный эксперимент с реальными радиолокационными снимками, подтвердивший адекватность и практическую ценность построенного распределения.

#### **Литература**

1. Carballo G.F., Fieguth P.W. Probabilistic cost functions for network flow phase unwrapping. // IEEE Transactions on Geoscience and Remote Sensing, vol. 38(5), pp. 2192–2201, 2000.
2. Chen C. W. Statistical-cost network-flow approaches to two-dimensional phase unwrapping for radar interferometry. // PhD thesis, Stanford University, 2001.
3. Goldstein R.M., Zebker H.A., Werner C.L. Satellite radar interferometry: Two-dimensional phase unwrapping. // Radio Sci., vol. 23, no. 4, pp. 713–720, July/Aug. 1988.
4. Guarnieri A. M. Using topography statistics to help phase unwrapping. //IEEE Proceedings online no. 20030565 (2003).

#### **Слова благодарности**

Автор благодарен к.т.н., зам. нач. отдела 00-30 ОАО "ВПК "НПО машиностроения" Елизаветину И.В. за плодотворное обсуждение работы и конструктивную критику.

# Численные методы и математическое моделирование

## Аппроксимация решения гиперсингулярного интегрального уравнения Прандтля тензорным поездом

*Андрюченко Назар Юрьевич*

*Студент (специалист)*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,*

*Россия*

*email: mail@gmail.com*

В работе рассматриваются вопросы численного решения гиперсингулярного уравнения Прандтля [1]:

$$-\int_0^1 \int_0^1 \frac{U(x, y)}{[(x-x_0)^2 + (y-y_0)^2]^{3/2}} dx dy = F(x_0, y_0) \equiv 4\pi V_z(x_0, y_0, 0) \quad (6)$$

Аналитических решений этого уравнения не известно, даже для достаточно „простых“ и „физических“ правых частей (например, для  $F(x_0, y_0) \equiv 1$ ).

Для дискретизации данного уравнения применяется метод дискретных вихрей [2], который приводит к системе линейных алгебраических уравнений относительно  $u_{ij}$ :

$$\sum_{i=1}^p \sum_{j=1}^p a_{ij}^{km} u_{ij} = F(x_0, y_0), \quad 1 \leq k, m \leq p \quad (7)$$

где

$$\begin{aligned}
 a_{ij}^{km} &= - \int_{x_{i-1}}^{x_i} \int_{y_{j-1}}^{y_j} \frac{dxdy}{[(x-x_{0k})^2 + (y-y_{0k})^2]^{3/2}} = \\
 &= - \frac{\sqrt{(x_{0k}-x_{i-1})^2 + (y_{0m}-y_j)^2}}{(x_{0k}-x_{i-1})(y_{0m}-y_j)} + \frac{\sqrt{(x_{0k}-x_i)^2 + (y_{0m}-y_j)^2}}{(x_{0k}-x_i)(y_{0m}-y_j)} - \\
 &\quad - \frac{\sqrt{(x_{0k}-x_i)^2 + (y_{0m}-y_{j-1})^2}}{(x_{0k}-x_i)(y_{0m}-y_{j-1})} + \frac{\sqrt{(x_{0k}-x_{i-1})^2 + (y_{0m}-y_{j-1})^2}}{(x_{0k}-x_{i-1})(y_{0m}-y_{j-1})}
 \end{aligned}$$

Размер полученной матрицы системы равен  $p^2 \times p^2$ , что, с учетом её плотности, очень много. Так, для сетки с  $p = 2^9$ , имеем матрицу  $2^{18} \times 2^{18}$ , на хранение которой уйдет порядка 0.5 терабайт памяти (не говоря о вычислительной сложности операций с этой матрицей). Возможным решением этой проблемы является поиск аппроксимации исходной матрицы  $A$  матрицей  $A_\varepsilon$ , которая близка к исходной ( $\|A - A_\varepsilon\| < \varepsilon$ ) и при этом определяется меньшим числом параметров. Один из возможных подходов к построению аппроксимаций такого сорта является аппроксимация тензорным поездом [3].

Под тензорным поездом понимается представление тензора (многомерного массива)  $B[i_1, i_2, \dots, i_d]$  в виде

$$B(i_1, i_2, \dots, i_d) = \sum_{\alpha_1, \dots, \alpha_{d-1}} G_1(i_1, \alpha_1) G_2(\alpha_1, i_2, \alpha_2) \dots G_{d-1}(\alpha_{d-2}, i_{d-1}, \alpha_{d-1}) G_d(\alpha_{d-1}, i_d)$$

Пусть  $G_1$  имеет размер  $n_1 \times r_1$ ,  $G_d - r_{d-1} \times n_d$  и остальные  $G_k - r_{k-1} \times n_k \times r_k$ . Введем обозначения:  $n = \max_k n_k$ ,  $r = \max_k r_k$ .

Для построения аппроксимации тензорным поездом матрицы  $A$  необходимо сначала сделать из матрицы тензор. Для этого  $i$  и  $j$  объявляются мультииндексами  $i = (i_1, \dots, i_d)$ ,  $j = (j_1, \dots, j_d)$  и производится перестановка индексов:  $A[i_1, \dots, i_d, j_1, \dots, j_d] = A[i_1, j_1, \dots, i_d, j_d]$ . Далее для тензора  $A[i_1, j_1, \dots, i_d, j_d]$  строится аппроксимация тензорным поездом.

В работе [3] приведен алгоритм ТТ-cross построения тензорного поезда по функции  $g$ , определяющей элементы тензора. Этот алгоритм требует всего лишь  $O(dnr^2)$  вычислений функции  $g$ . Тензорный поезд, построенный по данному алгоритму для матрицы системы (7) при  $p = 2^9$  на равномерной сетке (относительная погрешность аппроксимации во Фробениусовой норме порядка  $10^{-10}$ , время построения на обычном персональном компьютере менее 15 минут), занимает всего лишь 805 килобайт памяти, что более чем в  $10^5$  раз меньше, чем исходная матрица.

Матрицу системы можно заменить её аппроксимацией тензорным поездом и применять стандартные итерационные методы решения систем линейных алгебраических уравнений. В докладе приводится анализ скорости и точности полученных решений на равномерной и чебышевской прямоугольной сетках при использовании итерационных методов на подпространствах Крылова, а так же предлагается предобуславливатель для системы.

## Литература

1. Лифанов И.К. Метод сингулярных интегральных уравнений и численный эксперимент. М.: ТОО „Янус“. 1995.
2. Белоцерковский С.М., Лифанов И.К. Численные методы в сингулярных интегральных уравнениях. М.: Наука, 1985.
3. Оселедец И.В., Тыртышников Е.Е. Рекурсивное разложение многомерных тензоров. // Доклады РАН. 2009, Том 427, №1.

## Моделирование двумерных нестационарных газодинамических процессов методом

### ленточно-адаптивных сеток

*Белевский Владислав Викторович*

*Аспирант*

*Московский государственный технический университет имени Н.Э.*

*Баумана, Факультет фундаментальных наук, Москва, Россия*

*email: vachome@gmail.com*

В настоящей работе разработан метод решения задачи двумерной нестационарной газовой динамики, основанный на регулярных адаптивных сетках [1] с ленточной структурой, и, с его помощью, моделируются процессы нестационарного двумерного течения продуктов сгорания ракетного двигателя на твердом топливе (РДТТ).

Расчетная область изначально строится из блоков-примитивов в редакторе геометрии. Объединение локальных сеток осуществляется при помощи технологии ленточных адаптивных сеток [2-3]. Суть ее состоит во введении сплошной одноиндексной нумерации узлов и построении специального списка для такой нумерации.

Построение численного метода решения газодинамических уравнений производится с помощью модификации разностной схемы типа Мак-Кормака. Моделирование осуществляется для полного цикла работы РДТТ - от воспламенения основного заряда до полного его выгорания. Также, моделируется наличие мембраны в критическом сечении, разрыв которой происходит при достижении некоторого предельного значения давления. Для сокращения времени работы программы был разработан алгоритм решения задачи на нестационарных сетках с изменением числа узлов в процессе решения и применена технология параллельных вычислений для многопроцессорных (многоядерных) ЭВМ с общей памятью, как наиболее эффективная для данной задачи.

В результате данной работы разработано программное обеспечение, включающее редактор сеток, решатель и визуализатор с возможностью отображения видеоряда решения нестационарной задачи. При численном моделировании процессов получены распределения газодинамических параметров (давления, температуры газовой среды, вектора скорости движения газа).

## Литература

1. Годунов С.К., Забродин А.В., Иванов М.Я., Крайко А.Н., Прокопов Г.П. Численное решение многомерных задач газовой динамики. М.: Наука, 1976. – 400 с.
2. Димитриенко Ю.И., Изотова С.Г., Ануфриев С.Н., Захаров А.А. Численное моделирование трехмерных газодинамических процессов в камерах сгорания РДТТ на основе метода геометрически-адаптивных сеток // Вестник МГТУ им. Баумана. Сер. «Естественные науки». – 2005.-№3.
3. Димитриенко Ю.И., Изотова С.Г. Численное исследование нестационарных газодинамических процессов горения твердых топлив в камере РДТТ.-Аэрокосмические технологии: Труды Всероссийской научно-технической конференции/ Под ред. Р.П. Симоньянца. -М.:Изд-во МГТУ им.Н.Э.Баумана.-2003 .-С.144-154.

## Слова благодарности

Автор выражает признательность профессору, д.ф.-м.н. Димитриенко Ю.И. за помощь в подготовке тезисов.

## Численное решение трехмерной задачи двухфазного течения жидкостей в пористой среде

*Борисов Виктор Светославович*

*Аспирант*

*Якутский государственный университет им. М.К. Аммосова,  
Институт математики и информатики, Якутск, Россия  
email: refactorman@gmail.com*

Математическое моделирование реального процесса вытеснения из пористой среды одной жидкости другой жидкостью является достаточно научной проблемой. Оно основано на законах механики гетерогенных сплошных сред. Основные принципы и методы построения разностных методов решения краевых задач, в том числе и для математических моделей течения жидкостей и газов в пористых средах, разработаны в трудах А.Н.Тихонова, А.А.Самарского, Г.И.Марчука, П.Н.Вабищевича [2,5,6].

Целью работы является численное решение трехмерной задачи двухфазного течения жидкостей в пористой среде.

Широкое внедрение в промышленную практику даже достаточно вторичных методов разработки нефтяных месторождений, таких как вытеснение нефти водой приводит к необходимости углубленного изучения как самих математических моделей, так и методов их численной реализации [3].

Постановка задачи

$$\frac{\partial m \rho_{1s}}{\partial t} + \operatorname{div}(\rho_1, \vec{\omega}_1) = \sum_{i=1}^{N1} Q_i \delta(x - x_i, y - y_i, z - z_i), \quad (1)$$

$$(x, y, z) \in \Omega_1, t > 0$$

$$\frac{\partial m \rho_2 (1-s)}{\partial t} + \operatorname{div}(\rho_2, \vec{\omega}_2) = \sum_{i=1}^{N_2} Q_i \delta(x - x_i, y - y_i, z - z_i), \quad (2)$$

$$(x, y, z) \in \Omega_2, t > 0$$

$$\vec{\omega}_i = -k \frac{f_i}{\mu_i} \operatorname{grad}(p - \rho_i g z), \quad i = 1, 2 \quad (3)$$

$$m = \operatorname{const}, \rho_i = \operatorname{const}, s(x, y, z, 0) = s_0,$$

$$p(x, y, z, 0) = p_0, (x, y, z) \in \Omega$$

Задача решается с использованием метода фиктивных областей [2]. В работе дано решение актуальной научно-технической задачи, связанной с вытеснением нефти водой. На основе математического моделирования и численного эксперимента решена трехмерная задача двухфазного течения жидкостей в пористой среде. Предложен алгоритм и создана программа численного расчета, с использованием метода сопряженных градиентов.

#### Литература

1. Баренблатт. Г.И. Ентов В.М. Рыжик В.М. Движение жидкостей и газов в природных пластах. М.: Недра, 1984. – 211с.
2. Вабищевич П.Н. Метод фиктивных областей в задачах математической физики. –М.: Изд. МГУ. 1991. – 156 с.
3. Васильев В.И., Попов В.В., Тимофеева Т.С. Вычислительные методы в разработке месторождений нефти и газа. –Новосибирск: Изд-во СО РАН, 2000. –126 с.
4. Калиткин Н. Н. Численные методы, М.: Наука, 1978.– 512с.
5. Коновалов А.Н. Задача фильтрации многофазной несжимаемой жидкости. –Новосибирск: Наука, Сиб. отд-ние, 1988. –166 с.
6. Самарский А.А. Николаев Е.С. Методы решения сеточных уравнений. –М.: Наука, 1978. –562с.

#### Иллюстрации

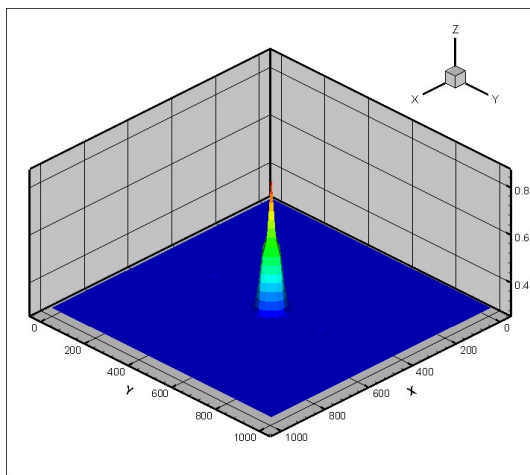


Рис. 1. График водонасыщенности, разрез по  $z=4$

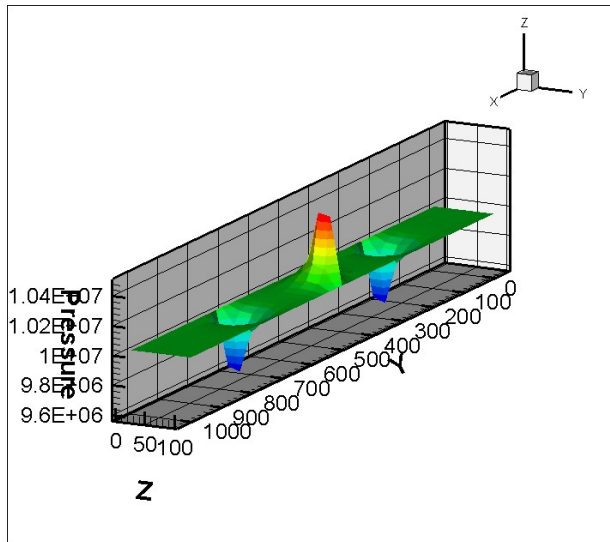


Рис. 2. График давления, разрез по  $x=25$



# Сплайн-интегрирование систем обыкновенных дифференциальных уравнений с запаздыванием

*Власов Сергей Александрович*

*Студент*

*Международный университет природы, общества и человека*

*"Дубна" филиал "Дмитров Направление Информатики и*

*Вычислительной Техники, Дмитров, Россия*

*email: vlasovs@inbox.ru*

В работе предложен новый подход численного интегрирования задачи Коши для систем обыкновенных дифференциальных уравнений (ОДУ) с запаздыванием. Он основан на построении шаг за шагом векторной сплайн-функции (для каждой компоненты решения), коэффициенты которой удовлетворяют:

- 1) условиям непрерывности и гладкости (с учетом специфики свойств решения исходной задачи);
- 2) исходной системе дифференциальных уравнений на очередном отрезке интегрирования (в одной точке или на подсетке Чебышева второго рода).

Для нахождения коэффициентов исходной сплайн-функции используется метод Ньютона. Для выбора текущей длины шага интегрирования предложен простой алгоритм, основанный на оценке текущей локальной погрешности по двум соседним сплайнам.

Получаемое решение является непрерывной и гладкой функцией, что позволяет использовать его в задачах оптимального выбора параметров динамической системы, возникающих, например, в задачах медицины и иммунологии [3].

Для проверки эффективности сплайн-интегрирования систем ОДУ с запаздыванием методом экстраполяции было построено решение известной задачи "хищник-жертва" с гарантированным порядком точности, которое использовалось при тестировании. Сравнение предлагаемого подхода проводилось с классическим методом, основанным на одношаговых методах Рунге-Кутты (3-го и 5-го порядка точности) [1,2] и построением интерполяционных многочленов Ньютона и Эрмита.

Выявлены преимущества и недостатки предлагаемого подхода.

## Литература

1. Хайрер Э., Нёрсетт С., Ваннер Г. Решение обыкновенных дифференциальных уравнений: Нежесткие задачи. М.: Мир, 1990.
2. Хайрер Э., Ваннер Г. Решение обыкновенных дифференциальных уравнений: Жесткие и дифференциально-алгебраические задачи. М.: Мир, 1999.
3. G. Bocharov, L. Cervantes-Barragan, R. Zust, K. Eriksson, V. Thiel, B. Ludewig «Mathematical modeling of the antiviral type I interferon response»: Proceeding of the FOSBE 2007: Stuttgart, Germany, September 9 – 12, 2007.

## Монотонный метод конечных объёмов для задачи конвекции-диффузии на сетках с многогранными ячейками

*Данилов Александр Анатольевич*

*Аспирант*

*Институт вычислительной математики РАН, -, Москва, Россия*

*email: a.a.danilov@gmail.com*

В работе предложен нелинейный метод конечных объёмов для стационарной задачи конвекции-диффузии. Диффузионный тензор в задаче может быть как полным, так и анизотропным, конвективная составляющая может быть преобладающей. Расчётная сетка состоит из произвольных выпуклых многогранников. Предлагаемый метод является монотонным, то есть он сохраняет неотрицательность решения. Монотонность метода достигается за счёт использования нелинейных двухточечных дискретизаций диффузионного и конвективного потоков через грани ячеек.

Трёхмерная версия метода[1] основана на двумерном варианте нелинейного метода конечных объёмов для многоугольных сеток[2]. В работе также представлены результаты экспериментальных расчётов, подтверждающих второй порядок сходимости предложенного метода.

### Литература

1. Danilov A., Vassilevski Yu. A monotone nonlinear finite volume method for diffusion equations on conformal polyhedral meshes // Russ.J.Numer.Anal.Math.Modelling. 2009. V.24(3). P.207-227.
2. Lipnikov K., Svyatskiy D., Vassilevski Yu. Interpolation-free monotone finite volume method for diffusion equations on polygonal meshes // J.Comp.Phys. 2009. V.228(3). P.703-716.

## Построение стохастических моделей в задачах среднесрочного прогнозирования солнечной активности

*Кудрявцев Ю.В., Шмалько Р.А.*

*Студент*

*Национальный технический университет Украины КПИ,  
Учебно-научный комплекс "Институт прикладного системного  
анализа Киев, Украина*

*email: kudriavcev@ukr.net*

Точные прогнозы уровня солнечной активности становятся все более важными, так как мы все более полагаемся на спутники, которые размещаются на околоземной орбите. Такие спутники зачастую выступают ключевым звеном в коммуникации, национальной обороне, а также являются источником важных научных наблюдений и исследований. Увеличение уровня солнечной активности может повлечь за собой существенные последствия в верхних слоях

атмосферы, оказывая влияние на орбиты спутников. Поэтому, долгосрочный и среднесрочный прогнозы уровня солнечной активности необходимы, чтобы помочь в разработке спутников, которые будут более долговременными и надежными.

Первый методом, предложенным для прогнозирования солнечных циклов, стал метод авторегрессии [2, 4] и с тех пор было применено большое разнообразие детерминированных методов прогнозирования. Большинство из них основано на определении корреляции между прошлым и будущим поведением процесса. Но такой подход дает довольно неточные результаты в связи с тем, что цикл солнечных пятен не является строго периодическим. За время наблюдений можно с уверенностью сказать, что от цикла к циклу изменяется как размер, так и продолжительность цикла. Известным примером является минимум Маундера. Все это указывает на стохастическое поведение солнечного цикла.

В данной работе для среднесрочного прогнозирования солнечной активности используется стохастическая модель солнечного цикла с априорно неизвестными статистическими характеристиками шумов. Стохастическая модель солнечного цикла отличается высокой степенью адекватности, так как детерминированной основой этой модели является аппроксимирующая кривая, построенная по историческим данным.

Проведен прогноз на 8, 12, и 18 месяцев вперед с помощью разработанного алгоритма среднесрочного прогнозирования солнечной активности, основанного на идентификации неизвестных статистических характеристик шумов предложенной модели и применении адаптивного фильтра Калмана [1]. Данные результаты сравнены с наиболее популярными подходами среднесрочного прогнозирования (Комбинированный метод [3], Стандартный метод [2]), которые применяются в мировом центре анализа солнечных данных SIDC). Введение стохастической составляющей в модель среднесрочного прогнозирования для различных прогнозов дает улучшение до 30

### Литература

1. Кудрявцев Ю.В., Шмалько Р.А., Подладчиков В.Н. Анализ эффективности методов прогнозирования на основе тестовых рядов Брауна // К.: Новые технологии, №3 2009. с.41–48.
2. Conway A.J. Time series, neural networks and the future of the Sun // *New Astronomy Reviews*. №42, 1988. p. 343–394.
3. Hanslemeier A., Denkmayr K., Weiss P. Longterm prediction if solar activity using the combined method // *Solar Phys.* №184, 1988. p. 213–218.
4. Hathaway D.H., Wilson R.M. A synthesis of solar cycle prediction techniques // *J. Geophys. Res.* №104, 1999. p. 375–388.

# Блочный тензорный метод типа сопряжённых градиентов для минимизации отношения Рэлея в QTT-формате

*Лебедева Ольга Сергеевна*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: lebedevaos@gmail.com*

Для задачи нахождения набора младших собственных векторов симметричной матрицы  $A\mathbf{X} = \mathbf{X}\Lambda$  (где матрица  $A = A^T$  размера  $n \times n$ ,  $\mathbf{X}$  — блок из  $d$  векторов, а  $\Lambda$  — диагональная матрица) существует ряд стандартных итерационных методов (см. книгу [3]). Они требуют не менее  $O(n)$  ячеек памяти и операций, что может быть велико для больших  $n$ . В ряде случаев эту проблему можно преодолеть, используя тензорные приближения. Одним из способов является QTT-разложение, описанное в [2]: если представить вектор длины  $n = 2^d$  в виде  $d$ -мерного тензора, и полученный тензор аппроксимировать в TT-формате [4], то количество параметров, необходимых для хранения полученного разложения, может составлять порядка  $O(\log_\alpha n)$  ячеек памяти. Кроме того, для векторов и матриц в таком формате существуют быстрые (с числом операций  $O(\log_\alpha n)$ ) и надёжные алгоритмы выполнения основных операций линейной алгебры [6].

Нами предложен новый блочный спектральный метод, в котором вместо всех векторов и матриц используются их приближения в QTT-формате. Номер вектора в блоке является дополнительной размерностью для тензора приближения, таким образом, приближения собственных векторов ищутся над общим адаптивным базисом. Мы стартуем с начального приближения в QTT-формате с небольшими рангами ядер, и проводим вычисления, аналогичные тем, что используются в стандартных спектральных алгоритмах. Сложность состоит в том, что у следующих приближений число параметров растёт. Поэтому мы определяем эффективную процедуру переаппроксимации векторов, которая будет проводиться на каждой итерации. Из-за переаппроксимации сходимость, вообще говоря, могла бы исчезнуть. В [1] нами было представлено доказательство сходимости модифицированных алгоритмов с аппроксимациями.

За основу нашего тензорного метода был выбран блочный метод сопряжённых градиентов [5]. Были реализованы несколько вариантов тензорного метода с «дожиманием» рангов: с жёсткими ограничениями на ранги ядер, с минимизацией рангов при ограничениях на вносимую ошибку и с ограничениями на ошибку и ранг. Новый метод был протестирован на примере конечно-разностных аппроксимаций эллиптических задач. Численные эксперименты подтвердили асимптотическую зависимость вида  $O(\log_\alpha n)$  времени работы метода от общего числа параметров задачи.

**Литература**

1. Лебедева О.С. «Блочный тензорный метод типа сопряжённых градиентов минимизации отношения Рэлея» ЖВМ и МФ, 2010, vol 50, № 5
2. Оселедец И.В. О приближении матриц логарифмическим числом параметров // ДАН, 2009. Т. 428:1. С.23-24.
3. Парлетт Б. Симметричная проблема собственных значений. – М.: Мир, 1983. – 384 с.
4. Oseledets I., Tyrtysnikov E. Breaking the curse of dimensionality, or how to use SVD in many dimensions // SIAM J. Sci. Comput., 2009. V. 31, №5. 3744-3759
5. D'yakonov E. G., Knyazev A. V. A group iteration method for finding the lowest eigenvalues. // Vestnik Moskov. Univ. Ser. XV Vychisl. Mat. Kibernet. №2, 1982. - p.29-34, 81.
6. TT-Toolbox <http://spring.inm.ras.ru/osel/>

## Математическое моделирование процесса заполнения лазерных мишеней ядерным топливом

*Малинина Елена Александровна*

*Студент*

*Московский государственный университет имени М.В. Ломоносова,*

*ВМК, Москва, Россия*

*email: malininal@yandex.ru*

Задача использования для практических целей энергии термоядерного синтеза, которая выделяется при слиянии легких ядер: дейтерия (D) и трития (T), D и D, D и  $^3\text{He}$  была сформулирована учеными полвека тому назад. Существует два основных подхода к решению поставленной задачи. Это магнитный термоядерный синтез и инерциальный термоядерный синтез (ИТС).

В ИТС используются лазерные мишени, которые, в частности, подвергаются прямому облучению лазером. Они представляют собой тонкостенные сферы, изготовленные из газопроницаемых материалов [1].

Изучаемая лазерная мишень представляет собой сферическую оболочку, состоящую из двух слоев. Внешний защитный шаровой слой изготовлен из полистирола или полиимида GDP. Внутренний слой пористый, изготовлен из пенополистирола. Как и в работах [2-4] была рассмотрена задача заполнения данной мишени газообразным ядерным топливом при различных толщинах слоев и коэффициентах газопроницаемости. Любой режим заполнения должен выдерживать разность между внешним и внутренним давлением на стенку сферы, не превышающую заданной величины, определяемой прочностными свойствами оболочки. Последнее требует знания зависимости от времени внутреннего давления, которое является функционалом от режима

нагнетания внешнего давления. Поскольку газ при высоком давлении не является идеальным, а прочностные параметры и коэффициенты газопроницаемости материалов стенки мишени могут быть не постоянными величинами, то на первое место в исследовании проблемы заполнения лазерных мишеней ядерным топливом выходят методы математического моделирования. Исследовалась модель для неидеального газа, поток которого через внешний слой подчиняется закону Фика, а в пористом слое описывается законом Дарси. Проведенный численный анализ модели показал, что в широком диапазоне практически важных значений коэффициента пористости и размера пор внутреннего пористого слоя – эти параметры мало влияют на время заполнения мишени газом, что особенно важно при выборе материала внутреннего слоя.

### Литература

1. Александрова И.В., Белоліпецкий А.А., Корешева Е.Р. Состояние проблемы криогенных топливных мишеней в современной программе инерциального термоядерного синтеза. // Вестник РАЕН, 2007, т.7, № 2, с. 37 – 42.
2. Белоліпецкий А.А.. Математическое моделирование и оптимизация процессов заполнения лазерных мишеней газом. // Вестник МГУ, серия 15 вычисл.матем. и кибернетика, 1997, № 4.
3. Aleksandrova I.V., Belolipetskiy A.A.. An efficient method for filling targets with a highly-pressurized gas fuel. // J. of the Moscow Physical Society, 1999, №9, p.325-335.
4. Aleksandrova I.V., Belolipetskiy A.A.. Mathematical models for filling polymer shells with a real gas fuel. // J. Laser and Particle Beams .1999, v.17, № 4, p.701-712.

### Слова благодарности

Выражаю глубокую признательность и благодарность моему научному руководителю Белоліпецкому Александру Алексеевичу за поддержку и помощь в разработке данной темы.

### Анализ применимости конечно-элементных моделей для решения нестационарных задач сейсродинамики и исследования эффективности применения горизонтальных сейсмических барьеров.

*Нафасов А.Э., Кузнецова М.Н.*

*Аспирант*

*Московский государственный строительный университет, Институт фундаментального образования, Москва, Россия*

*email: Aybek21@mail.ru*

Данная работа содержит в себе исследование численной устойчивости

конечно-элементных аппроксимаций, а также влияние количества и вида элементов разбиений на изменение перемещений на свободной границе упругой полуплоскости или упругого полупространства в задаче Лэмба. Решение этой задачи осуществлялось средствами программного комплекса ABACUS 6.8 .

Рассматривалась плоская и пространственная задачи Лэмба о распространении волн в упругой полуплоскости (полупространстве), вызванных действием гармонической сосредоточенной нагрузки, приложенной на свободной поверхности.

При анализе результатов расчета с гекса- и тетрагональными элементами для плоской и пространственной задач были обнаружены следующие результаты. Применение вида разбиения безразлично для плоской задачи, причем количество гексагональных элементов не имеет значения, а оптимальное число тетрагональных позволяет получить более точные результаты для амплитуд перемещений. Исследовать предел максимально возможного (без потери численной устойчивости) количества элементов в плоской задаче пока не получается из-за отсутствия достаточно мощной вычислительной техники. Для пространственной задачи применение тетрагональных элементов при количестве, не превышающем миллиона элементов, дает весьма точные результаты. Использование же гексагональных элементов в пространственной задаче количеством более 500 тысяч приводит к заметной потере численной устойчивости.

Анализ характеристик происшедших землетрясений, исследование повреждений зданий и сооружений, вызванных этими землетрясениями, а так же исследование природы сейсмических волн и закономерностей их распространения в различных грунтах и горных породах, позволили сделать вывод о том, что одним из оптимальных конструктивных решений по защите от землетрясений является горизонтальный барьер, позволяющий существенно уменьшать амплитуду сейсмических волн в защищаемой территории.

Горизонтальный барьер был смоделирован и проверен на взаимодействие с сейсмическими волнами с помощью метода конечных элементов в программе Abaqus 6.8. Задача решалась с помощью явной разностной схемы второго порядка точности.

Результаты расчета были представлены в виде соответствующих графиков, на основе которых можно сделать следующие утверждения:

- 1) При увеличении коэффициента Пуассона барьера от  $-0,75$  до  $0,5$  с шагом  $0,25$  и при коэффициенте Пуассона среды равным единице, магнитуда волн, прошедших барьер, практически не изменилась.
- 2) При увеличении модуля упругости барьера от  $0,1$  до  $20$  и при модуле упругости среды равным единице, магнитуда волн прошедших барьер уменьшается, но не значительно.
- 3) При увеличении плотности барьера от  $0,1$  до  $20$  и при плотности среды равной единице магнитуда волн, прошедших барьер, сильно уменьшается.

## Технология моделирования течений со свободной поверхностью в реалистичных сценах.

*Никитин Кирилл Дмитриевич*

*Аспирант*

*Институт вычислительной математики РАН, —, Москва, Россия*

*email: nikitink@dubki.ru*

Предлагается эффективная технология моделирования течения несжимаемой жидкости со свободной границей в топологически сложных областях [1,5].

Течение несжимаемой жидкости описывается уравнениями Навье-Стокса. Возможным компромиссом между сложностью, реалистичным поведением и контролем над потоком является технология, разрабатываемая в последние десятилетия группами специалистов, занимающихся компьютерной графикой [2,3,4]. В ее основе лежит проекционный метод, используемый для приближенного решения системы уравнений Навье-Стокса. Для работы со свободной поверхностью предлагается использовать два взаимодополняющих подхода: метод частиц и метод функции уровня [3,4].

Функция уровня задает положение свободной границы, а также позволяет получать локальную кривизну и расстояние до поверхности, которое нужно для визуализации. Решение уравнения Эйконала на каждом шаге позволяет поддерживать функцию уровня в виде расстояния до поверхности со знаком.

Для дискретизации уравнений Навье-Стокса предлагается использовать разнесенные гексаэдральные сетки, построенные по принципу восьмидерева, что позволяет динамически перестраивать сетку, сгущая ее к свободной поверхности в каждый момент времени. Кроме того использование гексаэдральных сеток дает оптимальное число степеней свободы, что также влияет на численную эффективность предлагаемой технологии.

Для взаимодействия жидкости с твердыми телами используется перенос свойств тела на сетку или принцип "жесткой жидкости"[2].

### Литература

1. Никитин К.Д. Технология расчета течений со свободной границей с использованием динамических гексаэдральных сеток. Численные методы, параллельные вычисления и информационные технологии. Изд-во МГУ, 2008. с. 183–198.
2. Carlson M. Rigid, Melting, and Flowing Fluid, PhD thesis, Georgia Institute of Technology, 2004.
3. Enright D., Fedkiw R., Ferziger J., Mitchell I. A hybrid particle level set method for improved interface capturing. J. Comp. Phys., 2002. V. 183. p. 83–116.
4. Fedkiw R., Osher S. Level Set Methods and Dynamic Implicit Surfaces. Springer-Verlag, 2002.



5. Nikitin K., Vassilevski Yu. Free surface flow modelling on dynamically refined hexahedral meshes. RJNAMM, V. 23. 2008. p. 469–485.

### Слова благодарности

Работа поддержана ФЦП "Научные и научно-педагогические кадры инновационной России" и грантом РФФИ 08-01-00159.

## Сравнительный анализ эффективности некоторых предобуславливателей.

*Никольский Илья Михайлович*

*Кандидат наук*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: haifly@rambler.ru*

Решение систем линейных алгебраических уравнений (СЛАУ) вида

$$Ax = b \tag{8}$$

является актуальной проблемой для численного решения многих задач науки и техники.

Разработано множество методов решения СЛАУ. В случае большой размерности матрицы  $A$  обычно применяются итерационные методы. К сожалению, очень часто итерации сходятся весьма медленно. Для ускорения сходимости используют так называемые предобуславливатели. Предобуславливатель является матрицей, которая в некотором смысле приближает матрицу  $A$ . Вместо исходной системы (1) решают эквивалентную ей предобусловленную систему

$$M^{-1}Ax = M^{-1}b. \tag{9}$$

Поскольку матрица  $M^{-1}A$  близка к единичной, итерационные методы при решении (2) сходятся лучше.

На сегодняшний день существует множество предобуславливателей различных классов. Количество итерационных методов решения СЛАУ также велико. Весьма важной и интересной задачей является нахождение оптимального сочетания метода и предобуславливателя для различных типов матриц. В данной работе с помощью вычислительного эксперимента проводится сравнительный анализ таких сочетаний.

### Литература

1. Николаев Е.С. Методы решения систем уравнений с разреженными матрицами. // Энциклопедия низкотемпературной плазмы. Сер. 5. Т. VII - 1. Математическое моделирование в низкотемпературной плазме. М.: Янус-К. 2009 г., с.345-548.

2. Y.Saad Iterative methods for sparse linear systems. SIAM, 2003.
3. Баландин М.Ю., Шурина Э.П. Методы решения СЛАУ большой размерности. Новосибирск: изд-во НГТУ. 2000

## Ньютоновские методы для задач оптимизации с комплементарными ограничениями

*Погосян Артур Леонович*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: pogosyan85@gmail.com*

Класс задач оптимизации с комплементарными ограничениями (ЗОКО) привлекает в настоящее время большое внимание специалистов. Многочисленные примеры возникновения таких задач в приложениях можно найти в [1] и цитированной там литературе. ЗОКО трудны для анализа и численного решения из-за неизбежной нерегулярности их ограничений.

В данной работе используется переформулировка ЗОКО, предложенная в [2]. Путем введения дополнительной переменной, т.е. «поднятия» исходной задачи, она сводится к задаче оптимизации с гладкими ограничениями-равенствами. В [2] использовалась переформулировка, обеспечивающая гладкость системы Лагранжа «поднятой» задачи. К сожалению, платой за эту гладкость является то, что при нарушении весьма обременительного условия строгой дополнительной нижнего уровня матрица Якоби этой системы в решении неизбежно оказывается вырожденной. В данной работе используется переформулировка, приводящая к негладкой системе Лагранжа. Однако, эта негладкость имеет весьма структурированный характер и допускает эффективное применение аппарата современного негладкого анализа и обобщенных версий метода Ньютона. В частности, обобщенные матрицы Якоби (элементы дифференциала Кларка) такой системы в естественных предположениях являются невырожденными. В сочетании со свойством полугладкости этой системы, указанное свойство гарантирует локальную сверхлинейную сходимость полугладкого метода Ньютона. Более того, квадрат невязки системы Лагранжа при этом оказывается гладкой функцией, что позволяет естественным образом глобализовать сходимость данного алгоритма, организовав на его основе метод спуска указанной гладкой оценочной функции. Однако, с практической точки зрения более перспективной представляется глобализация сходимости с использованием негладкой точной штрафной функции, что приводит к полугладкому методу последовательного квадратичного программирования.

### Литература

1. Luo Z.-Q., Pang J.-S., Ralph D. Mathematical programs with equilibrium constraints. Cambridge: Cambridge Univ. Press, 1996.

2. Stein O. Lifting mathematical programs with complementarity constraints //AOR-Preprint 2008. No. 3. University of Karlsruhe.

### Слова благодарности

Выражаю признательность своему научному руководителю профессору А. Ф. Измаилову.

## Алгоритм построения сильно различающихся решений некоторого класса обратных задач

*Сучков Егор Петрович*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: suchkov.egor@gmail.com*

Необходимость решения некорректно поставленных задач возникает во многих важных приложениях, например в проблеме управляемого термоядерного синтеза (УТС). При этом часто входные данные задачи определены лишь с некоторой погрешностью. Известно, что малым изменениям входных данных некорректно поставленной задачи могут соответствовать большие изменения её выходных данных [1]. Однако методы решения некорректных задач обычно разыскивают одно решение и не дают ответа на вопрос о существовании других существенно отличающихся от найденного решений, удовлетворяющим тем же входным данным в рамках заданной погрешности. Поэтому важным направлением является разработка методов построения всех сильно различающихся решений некорректных задач с неточно заданными входными данными.

Имея набор существенно различных решений обратной задачи можно по тем или иным дополнительным данным отобрать одно, соответствующее реальному физическому процессу.

Целью работы является изложение метода построения всех сильно различающихся

решений определённого класса некорректно поставленных задач с заданными с некоторой известной погрешностью входными данными.

Метод проиллюстрирован на примере решения одной из задач проблемы УТС - задачи реконструкции плотности тороидального тока в плазме по данным магнитной диагностики [2,3].

Результаты данной работы подтверждают вывод о том, что для некоторых обратных задач недостаточно найти какое-то решение, необходимо исследовать наличие других существенно отличающихся решений, также удовлетворяющих условиям задачи.

Предложенный в данной работе численный метод может быть использован для построения существенно различных решений или обоснования отсутствия

таких при решении достаточно широкого круга обратных задач математической физики.

### Литература

1. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. - М: Наука, 1986, 288 с.
2. Ю.Н. Днестровский, Д.П. Костомаров. Математическое моделирование плазмы. - М: Наука, первое издание 1982, 320 с., второе издание 1993, 336 с. (Издание на англ. языке: Yu.N. Dnestrovskij, D.P. Kostomarov. Numerical Simulations of Plasmas. Springer-Verlag, New-York, 1986.)
3. Ф.С. Зайцев. Математическое моделирование эволюции тороидальной плазмы. - Москва: МАКС Пресс, 2005, 524 с.

### Слова благодарности

Автор признателен члену-корреспонденту РАН Д.П. Костомарову и профессору Ф.С. Зайцеву за постановку задачи и обсуждение результатов исследований.

## Моделирование развития инфекции с использованием агентного подхода

*Улыбин Андрей Владимирович*

*Аспирант*

*Тамбовский государственный университет имени Г.Р. Державина,  
Институт математики, физики и информатики, Тамбов, Россия  
email: fatesmile@yandex.ru*

При моделировании инфекций и эпидемий с помощью аппарата обыкновенных дифференциальных уравнений не учитываются многие реальные свойства объектов [2]. Предполагается, что наиболее подходящим в данной области является имитационное моделирование с использованием агентного подхода [1]. В настоящем докладе представлена математическая модель имитационного моделирования динамики инфекции с использованием агентного подхода.

При моделировании использованы следующие допущения: модель развития инфекции представляет сложную систему, состоящую из множества взаимодействующих элементов (агентов); каждый агент имеет свойства, формируемые при его появлении (пол, возраст, максимальная продолжительность жизни, зараженность и т.д.); максимальная продолжительность жизни агента задается нормальным распределением вероятности при создании агента; агент может погибнуть раньше в случае поражения инфекцией, максимальная продолжительность жизни зараженного агента определяется в момент заражения; агенты формируются в системе в начальный момент времени, их параметры задаются распределениями вероятности.

Моделируемый объект состоит из конечного числа агентов. Каждый агент представляет собой объект класса с определенным набором свойств. Множество агентов включает в себя: инфицированные агенты; агенты, склонные к заражению; агенты, обладающие иммунитетом, агенты-носители инфекции. В модели определены правила взаимодействия агентов и изменения их свойств. Предполагается, что: инфицированный агент и агент-носитель инфекции посылают сигнал инфекции; при получении сигнала инфекции склонным к заражению агентом он меняет свое состояние на «заражен» и свои свойства в соответствии с правилами инфицирования; при получении сигнала инфекции агентом, обладающим иммунитетом, он с большой вероятностью не изменяет своих свойств. Аналогичные правила создаются для иных взаимодействий.

Проведенные исследования над моделью на примере ВИЧ инфекции, показали, что характер динамики ВИЧ инфекции по реальным данным и по модели совпадают. Выявлено, что общее количество ВИЧ инфицированных гораздо больше официально зарегистрированных случаев. Это предположение подтверждается большинством экспертных оценок, а также данными Объединенной программы ООН по ВИЧ /СПИДу [3]. Таким образом, модель вполне можно считать пригодной для практических целей при изучении инфекции.

### Литература

1. Арзамасцев А.А., Соломина О.А. Моделирование роста биологической популяции на плоскости // Математическое моделирование. 2009. No. 4. С. 59-64.
2. Perelson, A. S. Modelling viral and immune system dynamics. // Nature Reviews Immunology. 2002. No. 1. С. 28-36.
3. ЮНЭЙДС.ру: <http://www.unaids.ru>

### Вычисления из первых принципов пути реакции изомеризации в молекулярной системе методом метадинамики.

*Шумкин Георгий Николаевич*

*Аспирант*

*Московский государственный университет имени М.В. Ломоносова,  
Факультет вычислительной математики и кибернетики, Москва,  
Россия*

*email: georgiy-sh@yandex.ru*

В работе представлены результаты математического моделирования "из первых принципов" молекулярного переключения на основе индуцированной током сканирующего туннельного микроскопа реакции изомеризации в молекуле нафталицианина [4]. Вдоль введенной координационной переменной

вычислен путь химической реакции, и найдена высота энергетического барьера, который следует преодолеть для прохождения реакции водородной изомеризации. Поверхность свободной энергии Гиббса и путь реакции проанализированы методом "метадинамики" [3] в рамках численного квантовомеханического кода молекулярной динамики Кар-Парринелло "CPMD" [1-2]. Расчеты проведены на супер-ЭВМ IBM Blue Gene/P, установленной на факультете ВМК МГУ.

### Литература

1. Andreoni W., Curioni A. New Advances in Chemistry and Materials Science with CPMD and Parallel Computing // Parallel Computing. 2000. 26. P.819-842.
2. Car R., Parrinello M. Unified approach for molecular dynamics and density-functional theory // Phys. Rev. Lett. 1985. 55. P.2471-2474.
3. Ensing B., Laio A., Parrinello M. and et. al. A Recipe for the Computation of the Free Energy Barrier and the Lowest Free Energy Path of Concerted Reactions // J. Phys. Chem. B. 2005. 109. N 14. P.6676-6687.
4. Liljeroth P., Repp J., Meyer G. Current-Induced Hydrogen Tautomerization and Conductance Switching of Naphthalocyanine Molecules // Science. 2007. 317. P.1203-1206.