

Федеральный исследовательский центр  
«Информатика и управление»  
Российской академии наук

Труды Второй молодежной  
научной конференции  
«Задачи современной информатики»



2015

Федеральный исследовательский центр  
«Информатика и управление»  
Российской академии наук

Труды Второй молодежной  
научной конференции  
«Задачи современной информатики»

Москва

2015

*Печатается по решению Ученого совета  
Федерального исследовательского центра  
«Информатика и управление»  
Российской академии наук*

Рецензент: доктор физико-математических наук,  
профессор Шоргин С. Я.

Труды Второй молодежной научной конференции  
«Задачи современной информатики» – М.: ФИЦ ИУ РАН,  
2015. – 319 с. – ISBN 978-5-91993-054-9.

Сборник трудов содержит 45 научных докладов, отобранных для публикации в рамках Второй молодежной научной конференции «Задачи современной информатики» (ЗСИ-2015) 29–30 октября 2015 года. Конференция организована Федеральным исследовательским центром «Информатика и управление» Российской академии наук (г. Москва) совместно с факультетом Вычислительной математики и кибернетики МГУ им. М.В. Ломоносова и Российским университетом дружбы народов.

**Вторая молодежная научная конференция «Задачи современной информатики» проводится при финансовой поддержке Российского фонда фундаментальных исследований (проект №15-37-10461).**

# Содержание

	Введение.....	8
<i>Агафонов Е.С., Белюсов В.В., Горшенин А.К.</i>	Опыт разработки стохастических информационных технологий для поддержки научных исследований в экономике.....	11
<i>Андрейчук А.А., Яковлев К.С.</i>	Планирование траекторий для коалиции агентов. Алгоритм MAPR.....	18
<i>Бегишев В.О.</i>	Математическая модель для анализа характеристик интерференции при взаимодействии устройств в беспроводной сети.....	25
<i>Бесчастный В.А.</i>	Оценка качества аппроксимации хардкор процессов Матерна пространственными пуассоновскими процессами.....	32
<i>Боковой А.В.</i>	Метод одновременного картирования и локализации по видеопотоку, полученному с единственной камеры, LBA-vSLAM.....	39
<i>Болдырева А.В.</i>	Применение метода МГУА на основе интенсивности поисковых запросов в сети Интернет для прогноза рынка недвижимости.....	46
<i>Борисов А.И.</i>	Оценка многомерной плотности распределения.....	52
<i>Васильев И.Ю., Гайдамака Ю.В.</i>	Имитационная модель буферизации видеопотока в одноранговой сети с учетом геолокации и активности пользователей.....	58
<i>Гиоргис Э.М.</i>	Анализ естественных и искусственных изображений в рекламе посредством непараметрической	

	оценки функции распределения плотности пикселей.....	65
<i>Горбунов П.В.</i>	Опыт разработки и испытаний флуориметра для экспресс-оценки жизнеспособности бактериальных культур.....	72
<i>Горбунова А.В.</i>	Оценка времени отклика обработки запросов в системе облачных вы- числений.....	79
<i>Гусынин В.И.</i>	Алгоритмы поиска арбитражных возможностей на валютном рынке..	86
<i>Зайцев Р.Д., Бритков В.Б.</i>	Применение языка R для много- мерной кластеризации временных рядов с целью анализа динамики научно-технического развития.....	92
<i>Лесько С.А., Жуков Д.О.</i>	Описание распространения инфор- мации в сетях социальных связей на основе теории перколяции.....	99
<i>Лесько С.А., Жуков Д.О.</i>	Структура информационного про- странства: директор и самоподобие.	106
<i>Листопад С.В.</i>	Концептуальная модель коллекти- вов интеллектуальных агентов для решения сложных задач.....	113
<i>Неменко М.В., Лесько С.А., Жуков Д.О.</i>	Балансировка потоков в загружен- ных транспортных сетях на основе стохастической модели.....	121
<i>Неменко М.В., Лесько С.А., Жуков Д.О.</i>	Стохастическая динамика перехо- дов между состояниями в сетях со- циальных связей, самоорганизация и скачки настроений.....	128
<i>Петрушкина О.А., Нуриев В.А., Спиридонова Е.С.</i>	Алгоритм изучения категории не- определенности в русском языке на основе корпусных данных.....	136

<i>Попкова Н.А., Инькова О.Ю., Зацман И.М., Кружков М.Г.</i>	Методика построения моноэквивалентностей в надкорпусной базе данных коннекторов.....	143
<i>Решетняк Н.В., Постников В.В.</i>	Обнаружение границ препятствий в видеопотоке, полученном с фронтальной стереопары.....	154
<i>Ровенская О.И.</i>	Исследование микротечений газа с помощью гибридного метода.....	160
<i>Румовская С.Б.</i>	Методы и средства информатики для диагностики артериальной гипертензии в лечебно-профилактических учреждениях широкого профиля.....	168
<i>Семенова О.В., Власкина А.С., Норкина М.А., Шевякова К.А.</i>	Модель бизнес-процесса в виде сети массового обслуживания.....	175
<i>Сеченых П.А., Абгарян К.К.</i>	Реляционная модель хранения данных информационной поддержки задач структурного материаловедения.....	181
<i>Сидоркин И.И.</i>	Метод детектирования машинопечатных текстовых информационных зон отпечатков документов строгого учета.....	187
<i>Сорокин А.И.</i>	Метод фильтрации сегментированных изображений.....	194
<i>Степанов Б.А., Дьяченко Ю.Г., Рождественский Ю.В.</i>	Самосинхронный умножитель для распределенных вычислений.....	201
<i>Федуро А.А., Филипова В.Р., Маркова Е.В., Мокров Е.В., Гудкова И.А.</i>	Модель для анализа вероятности прерывания обслуживания в сети 3GPP LTE с системой LSA.....	208

<i>Флёрова А.Ю.</i>	Исследование возможностей прогнозирования нефтедобычи с помощью математического моделирования.....	215
<i>Хилько Д.В., Шикунов Ю.И., Степченков Ю.А.</i>	Особенности программной реализации имитационной модели потоковой рекуррентной архитектуры.....	220
<i>Чернецов А.М.</i>	Возможности параллельного программирования в пакетах MATLAB и Maple.....	228
<i>Шинов А.А.</i>	О методах оценки цветопередачи источников света путём анализа спектра с использованием языка программирования Python.....	235
<i>Шмалько Е.Ю., Дивеев А.И.</i>	Метод многослойного сетевого оператора в задаче синтеза системы управления группой роботов.....	241
<i>Этезов Ш.А.</i>	Анализ системы поллинга с пороговым управлением.....	248
<i>Яшина В.В.</i>	Современное состояние алгебраической теории анализа изображений. Обзор.....	254
<i>Бурашникова А.И.</i>	Интеграция многокритериальных методов с целью выявления предпочтений.....	261
<i>Дорофеева А.В., Королев В.Ю.</i>	Оценки скорости сходимости в центральных предельных теоремах для случайных сумм при ослабленных моментных условиях.....	269
<i>Дубнов Ю.А.</i>	Оценка параметров линейной регрессии методом максимальной энтропии.....	276
<i>Заспа А.Ю.</i>	Исследование среднеквадратичного риска при обработке звуковых сигналов FDR-порогом.....	283

<i>Ковальчуков Р.Н., Гайдамака Ю.В.</i>	Имитационная модель для оценки интерференции в одноранговых се- тях беспроводных устройств с пря- моугольными кластерами.....	290
<i>Коротышева А.В., Киселева К.М., Сатин Я.А.</i>	Эргодичность и устойчивость си- стемы обслуживания с одним серве- ром.....	297
<i>Мазурин Д.С.</i>	Динамическое распределение транспортных потоков в сети круп- ного города.....	303
<i>Мейханаджян Л.А., Зарядов И.С., Разумчик Р.В.</i>	О стационарных характеристиках системы Geo/Geo/1/r с полным об- новлением, зависящим от состоя- ния очереди.....	309
<i>Пономаренко- Тимофеев А.А.</i>	Имитационная модель сценария совместного использования частот аэропортом и оператором сети 3GPP LTE.....	314

# Введение

В 2015 году Вторая молодежная научная конференция «Задачи современной информатики» (ЗСИ-2015) проводится Федеральным исследовательским центром «Информатика и управление» Российской академии наук (ФИЦ ИУ РАН) совместно с факультетом Вычислительной математики и кибернетики МГУ им. М.В. Ломоносова и Российским университетом дружбы народов при финансовой поддержке Российского фонда фундаментальных исследований (проект №15-37-10461).

К рассмотрению принимались статьи на русском языке от молодых исследователей, возраст которых не превышает 35 лет (включительно) для кандидатов наук, а также в случае отсутствия ученой степени, и 39 лет (включительно) для докторов наук. Были предложены следующие основные направления научных исследований и разработок в области информатики:

- теоретические основы информатики и информационных технологий;
- математические модели и методы исследования информационно-телекоммуникационных и стохастических систем;
- информационные технологии накопления, хранения, поиска, обработки, преобразования, отображения, защиты и передачи информации;
- архитектура, математическое и программное обеспечение вычислительных систем и сетей;
- параллельные и распределенные высокопроизводительные вычисления, потоковая обработка сверхбольших массивов данных;
- математические методы и информационные технологии системного анализа и управления;
- искусственный интеллект, когнитивные исследования;

- методы исследования и алгоритмы решения задач вычислительной математики, математической физики, теории оптимизации.

В состав *Программного комитета* вошли: д.ф.-м.н., профессор *Королев В. Ю.* (ВМК МГУ) – председатель; д.т.н., профессор *Самуйлов К. Е.* (РУДН) – заместитель председателя; к.ф.-м.н., доцент *Гуревич И. Б.* (ФИЦ ИУ РАН (ВЦ им. А. А. Дородницына РАН)); д.ф.-м.н., профессор *Осипов Г. С.* (ФИЦ ИУ РАН); д.т.н., профессор *Синицын И. Н.* (ФИЦ ИУ РАН (ИПИ РАН)); д.ф.-м.н., доцент *Шестаков О. В.* (ВМК МГУ).

В состав *Организационного комитета* вошли: д.ф.-м.н., профессор *Шоргин С. Я.* (ФИЦ ИУ РАН) – председатель; к.ф.-м.н. *Горшенин А. К.* (ФИЦ ИУ РАН (ИПИ РАН)) – заместитель председателя; к.т.н. *Белоусов В. В.* (ФИЦ ИУ РАН (ИПИ РАН)); к.т.н. *Бульчев А. В.* (ФИЦ ИУ РАН (ИСА РАН)); к.ф.-м.н., доцент *Гайдамака Ю. В.* (РУДН); к.т.н. *Тарханов И. А.* (ФИЦ ИУ РАН (ИСА РАН)); к.т.н. *Трусова Ю. О.* (ФИЦ ИУ РАН (ВЦ им. А. А. Дородницына РАН)); к.ф.-м.н. *Яковлев К. С.* (ФИЦ ИУ РАН (ИСА РАН)).

Для передачи опыта молодым учёным в рамках конференции ЗСИ-2015 были представлены следующие пленарные доклады:

- «*Распознавание образов: модели процедур обработки данных*» (докладчики – академик РАН *Журавлев Ю. И.*, чл.-корр. РАН *Рудаков К. В.*);
- «*Современные научно-практические аспекты развития информационных технологий*» (докладчик – д.т.н., профессор *Зацаринный А. А.*);
- «*Метод GFS на базе простых испытаний Монте Карло для решения задач глобальной оптимизации (эксперименты, приложения)*» (докладчик – чл.-корр. РАН *Попков Ю. С.*).

В конференции приняли участие студенты, аспиранты и молодые учёные из различных российских научных

центров и высших учебных заведений, представлявшие:

- Федеральный исследовательский центр «Информатика и управление» Российской академии наук, включая Институт проблем информатики РАН ФИЦ ИУ РАН, Вычислительный центр им. А. А. Дородницына РАН ФИЦ ИУ РАН, Институт системного анализа РАН ФИЦ ИУ РАН, Калининградский и Орловский филиалы ФИЦ ИУ РАН;
- факультет Вычислительной математики и кибернетики МГУ им. М.В. Ломоносова;
- Российский университет дружбы народов;
- Московский физико-технический институт;
- Московский государственный университет информационных технологий, радиотехники и электроники;
- Вологодский государственный университет;
- Московский авиационный институт;
- Национальный исследовательский технологический университет «МИСиС»;
- Московский педагогический государственный университет;
- Московский технический университет связи и информатики;
- Высшую школу экономики;
- Российскую академию народного хозяйства и государственной службы при Президенте Российской Федерации.

Всего было подано 60 докладов, из них в процессе рецензирования для издания были отобраны 45 работ. Настоящий сборник содержит труды конференции, публикуемые, преимущественно, в авторской редакции (за исключением внесения дополнительного форматирования и исправления ряда орфографических и пунктуационных ошибок). Все рисунки напечатаны в градациях серого цвета независимо от первоначального авторского варианта.

# Опыт разработки стохастических информационных технологий для поддержки научных исследований в экономике

Е.С. Агафонов<sup>1</sup>, В.В. Белоусов<sup>2</sup>, А.К. Горшенин<sup>3</sup>

**Аннотация:** Работа посвящена стохастическим информационным технологиям поддержки научных исследований в экономике. Дана классификация информационных технологий моделирования стохастических систем. Представлены результаты работы двух программных продуктов. Первый инструментальный программный пакет предназначен для оптимизации жизненного цикла изделий наукоемкой продукции, второй - для управления запасом некоторого продукта.

**Ключевые слова:** информационные технологии, стохастические информационные технологии, пакет прикладных программ, стохастические системы, изделие наукоемкой продукции, жизненный цикл, управление запасом, оптимизация.

## Введение

Решение большинства современных научных задач невозможно без применения информационных технологий (ИТ). К характерным функциям ИТ поддержки научных исследований в экономике можно отнести следующие: интенсивная обработка больших объемов информации, оперативная реакция на структурные изменения моделей, учет случайных факторов разной природы, работа с базами данных, визуализация информации. Важную роль при решении данных задач играют стохастические ИТ, моделирующие процессы в стохастических

---

<sup>1</sup> ФИЦ ИУ РАН (ИПИ РАН); 1983; eagafonov@ipiran.ru

<sup>2</sup> ФИЦ ИУ РАН (ИПИ РАН); 1977; к.т.н.; vbelousov@ipiran.ru

<sup>3</sup> ФИЦ ИУ РАН (ИПИ РАН); 1986; к.ф.-м.н.; agorshenin@ipiran.ru

системах (СтС).

Современное понятие «обработки» информационных процессов с помощью ЭВМ предполагает взаимозависимое в рамках единой технологии решение комплекса методологических, алгоритмических и технических вопросов, обеспечивающих возможность использования информационных процессов для решения разнообразных практических задач. Такое понимание «обработки» с помощью ЭВМ позволяет выделить общие аспекты решения задач, связанных с обработкой информационных процессов, и более четко сформулировать понятие информационных технологий моделирования систем как эффективных средств решения задач научных исследований.

## **1. Классификация ИТ моделирования СтС**

На практике обычно создается ИТ, рассчитанная на выполнение с ее помощью некоторой основной функции, что связано с необходимостью решения нескольких типовых задач статистических исследований. Перечень основных функций довольно ограничен, с другой стороны, выполнение этих функций может потребоваться во многих применениях. Это делает целесообразным выделение функционально-ориентированных, предметно-ориентированных и проблемно-ориентированных ИТ:

- функционально-ориентированные ИТ предназначены для реализации типовых, относительно автономных, задач исследования СтС, они обычно обладают довольно высокой степенью универсальности;
- предметно-ориентированные ИТ предназначены для решения специфической задачи статистических исследований в конкретной области, такие ИТ максимальным образом удовлетворяют частным требованиям данного применения и могут обладать наименьшей степенью универсальности;
- проблемно-ориентированные ИТ занимают про-

межучточное положение между функционально-ориентированными и предметно-ориентированными ИТ, они реализуют некоторые общие прикладные проблемы исследований СтС.

Можно сформулировать два основных подхода к созданию ИТ анализа СтС:

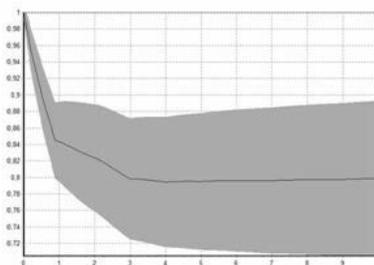
- создание библиотек подпрограмм для различных языков программирования, например, таких как C#, C++, Fortran, Pascal и др.;
- создание самостоятельных диалоговых пакетов.

Последние, в свою очередь, можно классифицировать по степени полноты пакета, разделив на полные, которые содержат все необходимые инструментальные средства и не требующие расширения, и неполные пакеты, расширяющие возможности некоторой базовой системы компьютерной математики (СКМ), например, MATLAB, Mathematica, Maple, Mathcad, Derive, VisSim и др. СКМ могут успешно применяться при решении многих задач математического моделирования. При этом они реализуют как численные, так и аналитические методы решения задач. В технике проведения научных расчетов СКМ стали своего рода суперкалькуляторами, практически мгновенно решающими сложные уравнения, вычисляющие интегралы и производные, строящими графики любых функций и т.д.

## **2. ИТ моделирования жизненного цикла изделий наукоемкой продукции**

Использование информационно - вычислительных систем в сфере управления стоимостью жизненного цикла изделий наукоемкой продукции (ЖЦ ИНП) основано на CALS - методологии (Continuous Acquisition and Life cycle Support - непрерывной информационной поддержке поставок и жизненного цикла). Описание данной концепции подробно представлено в [1–3]. Ее основу составляют технологии логистической поддержки (ЛП), средства и процедуры которой должны обеспечи-

вать минимизацию затрат в системе послепродажного обслуживания (СППО). Однако при всем совершенстве и многообразии предлагаемых в этой области технологий, фактически остается нерешенной главная проблема – проблема оптимального проектирования системы обслуживания, ремонта и снабжения для обеспечения эксплуатации изделий с минимальной стоимостью на годы вперед, вплоть до их списания. В работах ИПИ ФИЦ ИУ РАН (подразделение Институт проблем информатики Федерального исследовательского центра «Информатика и управление» РАН) развиваются новые научные стохастические подходы к решению сформулированной проблемы. Подробное, систематичное изложение данных подходов можно найти в [1,2,4,5]. В данном разделе статьи демонстрируются результаты работы полного пакета «Оптимизация СППО», созданного на основе предлагаемых подходов. Решается упрощенная задача оптимизации – минимизация затрат заказчика на поставки запасных частей ИПП на годы вперед, вплоть до списания. На Рис. 1 представлен долгосрочный прогноз (10 лет) одной из основных характеристик технико-экономической системы – коэффициента исправности с указанием доверительного интервала. На Рис. 2 представлена финансовая программа на 10 лет, сформированная на основе рассчитанного программой оптимального плана закупок запасных частей ИПП по годам.



**Рис. 1.** Прогноз коэффициента исправности

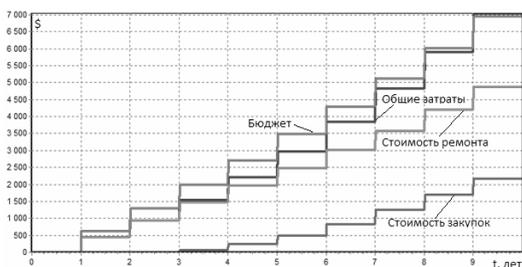
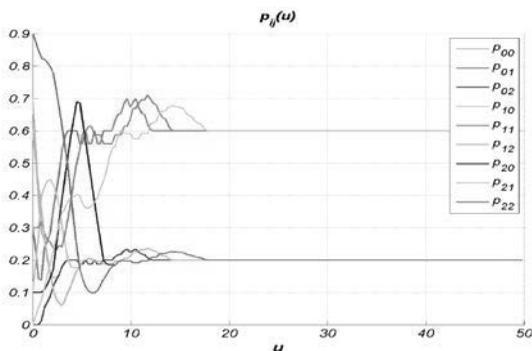


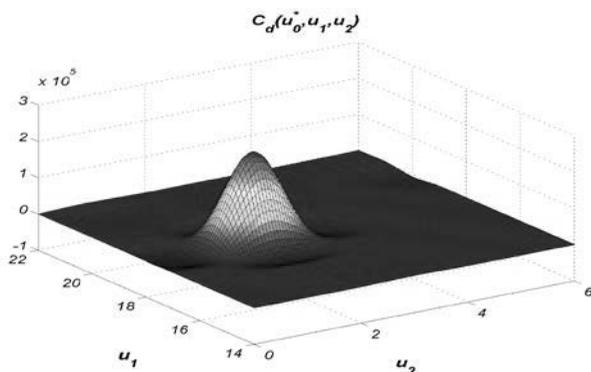
Рис. 2. Программа закупок и финансовый запрос

### 3. ИТ управления запасами продуктов

В данном разделе статьи демонстрируется работа программного пакета «Inventory», созданного на базе системы MATLAB. В пакете реализована дискретная стохастическая модель управления запасом некоторого продукта, основанная на использовании управляемого полумарковского процесса [6]. На первом этапе работы программы вводятся вероятностные характеристики полумарковского процесса и стоимостных функционалов, затем выполняются аналитические преобразования для получения в явном виде функционалов, характеризующих качество управления. На втором этапе работы программы решается задача поиска экстремума заданной функции нескольких вещественных переменных - оптимального управления. На Рис. 3 показаны промежуточные результаты расчета – вероятностные характеристики полумарковского процесса. На Рис. 4 представлен графический результат поиска экстремума функции оптимального управления.



**Рис. 3.** Вероятностные характеристики полумарковского процесса



**Рис. 4.** Функция оптимального управления

## Выводы

На основе разработанного методического обеспечения, коллективом сотрудников ФИЦ ИУ РАН создано инструментальное программное обеспечение для моделирования стохастических систем.

Полный пакет «Оптимизация СППО», написанный на языке Pascal, осуществляет автоматический поиск оптимальных программ поставок СЧИ и объемов ремонта по годам в условиях финансово-бюджетных ограничений, что позволит эффективно решать задачу оптимизации для сложных гибридных

систем послепродажного обслуживания.

Неполный пакет «Inventory», созданный в среде MATLAB, используется для моделирования полумарковской модели управления произвольными запасами (продуктами). Он позволяет находить оптимальное управление через поиск экстремума функции нескольких переменных.

## Список литературы

1. *Синицын И. Н., Шаламов А. С.* Лекции по теории систем интегрированной логистической поддержки. – М.: ТОРУС ПРЕСС, 2012. – 624с.
2. *Пугачев В. С., Синицын И. Н.* Теория стохастических систем. – М.: Изд-во «Логос», 2000 и 2003. (1 и 2 изд.) [Англ. пер.: *Stochastic Systems. Theory and Applications.* – Singapore, World Scientific, 2001].
3. *Киркпатрик Д.* Основные проблемы, связанные с определением стоимости жизненного цикла // Бюл. техн. информации по мат. междунар. конф. «Стоимость жизненного цикла изделий оборонной и авиационно-космической техники», 2000. С. 149–155.
4. *Синицын И. Н., Шаламов А. С.* Методологические аспекты современной интегрированной логистической поддержки изделий наукоемкой продукции // Системы высокой доступности, 2011. Т. 7. №4. С. 48–74.
5. *Синицын И. Н., Шаламов А. С., Сергеев И. В., Синицын В. И., Корепанов Э. Р., Белоусов В. В., Агафонов Е. С., Шоргин В. С.* Методы и средств анализа и моделирования стохастических систем интегрированной логистической поддержки // Системы и средства информатики. – М.: ИПИ РАН, 2012. Вып. 22. №2. С. 3–28.
6. *Gorshenin A. K., Belousov V. V., Shnourkoff P. V., Ivanov A. V.* Numerical Research of the Optimal Control Problem in the Semi-Markov Inventory Model // AIP Conference Proceedings, 2015. Vol. 1648. P. 250007-1–250007-4.

# Планирование траекторий для коалиции агентов. Алгоритм MAPP\*

А.А. Андрейчук<sup>4</sup>, К.С. Яковлев<sup>5</sup>

**Аннотация:** В работе рассматривается задача планирования совокупности траекторий для коалиции агентов, проводится анализ существующих методов, решающих указанную задачу, описывается алгоритм MAPP, приводятся результаты его экспериментальных исследований.

**Ключевые слова:** планирование траектории, поиск пути, мультиагентное планирование, MAPP

## Введение

Планирование совокупности траекторий для коалиции агентов является на порядок более сложной проблемой, чем планирование траектории для отдельного агента, и относится к классу сложности PSPACE[1], поэтому разработка вычислительно эффективных алгоритмов её решения является весьма актуальной научно-технической задачей.

Традиционно задача планирования траектории рассматривается как задача поиска пути на графе, вершинам которого соответствуют положения агента в пространстве, а ребрам – возможные переходы между ними. Весьма распространено использование графа специального вида – метрического топологического графа (в англоязычной терминологии – grid) [2].

Метрический топологический граф (МТ-граф) – взвешенный неориентированный граф, каждый элемент (вершина)

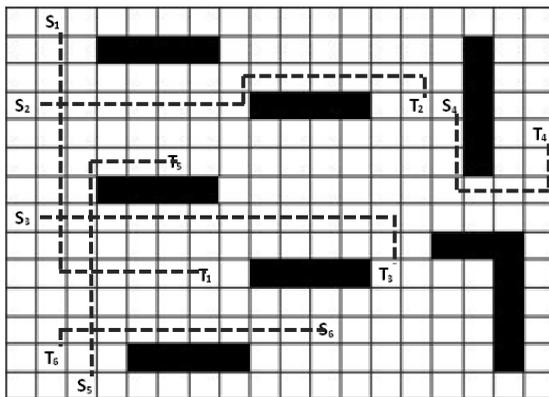
---

\*Работа выполнена поддержке РФФИ в рамках научного проекта № 15-37-20893.

<sup>4</sup> РУДН; 1993 г.р.; andreychuk@mail.com

<sup>5</sup> ФИЦ ИУ РАН (ИСА РАН); 1983 г.р.; к.ф.-м.н.; yakovlev@isa.ru

которого соответствует некоторой области на плоскости. Элементы МТ-графа могут быть либо проходимы, либо непроходимы для агента. Ребрами являются пары смежных проходимых вершин. Графически МТ-граф задачи планирования можно представить в виде таблицы (см. Рис. 1). Светлым клеткам на рисунке соответствуют проходимые вершины, темным – непроходимые, пунктирными линиями обозначены пути для каждого из агентов коалиции.



**Рис. 1.** Графическое представление МТ-графа

Задано непустое множество агентов  $U$  и для каждого агента из  $u \in U$  задана пара вершин старт–цель  $(s_u, t_u)$ ,  $s_u \neq s_v \wedge t_u \neq t_v \forall u, v \in U$ . Путь из  $s_u$  в  $t_u$  для агента  $u$  – это последовательность смежных вершин:  $\pi(u) = (l_0^u, l_1^u, \dots, l_{k_u}^u)$ ,  $l_0^u = s_u, l_{k_u}^u = t_u$ . В начальный момент времени агенты располагаются в соответствующих начальных вершинах. За один такт каждый из агентов может переместиться из вершины  $l_i^u$  в вершину  $l_{i+1}^u$  либо остаться в вершине  $l_i^u$ . Множество моментов времени (тактов), в которые агент  $u$  занимает вершину  $l_i^u$  обозначим как  $time(l_i^u)$ . Задача планирования состоит в построении множества неконфликтных путей, т.е. таких путей, что выполняется условие  $\forall u, v \in U, \forall l_i^u = l_j^v \text{ time}(l_i^u) \cap \text{time}(l_j^v) = \emptyset, i \in [0, k_u], j \in [0, k_v]$ .

# 1. Методы решения

Алгоритмы планирования траекторий для совокупности агентов можно разделить на две группы. К первой группе относятся алгоритмы, принцип работы которых основывается на централизованном подходе к решению задачи, то есть при планировании учитываются положения (в каждый момент времени) всех агентов [3,4]. Использование подобного подхода позволяет обеспечить полноту соответствующих алгоритмов, т.е. гарантировать нахождение решения (если оно существует). Однако, из-за низкой вычислительной эффективности эти алгоритмы не применимы в практических задачах, когда число агентов велико.

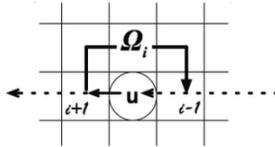
Алгоритмы второй группы используют децентрализованный подход. Принцип работы этих алгоритмов заключается в независимом (с некоторыми ограничениями) планировании траектории для каждого из агентов, с последующим согласованием траекторий. К этой группе относятся алгоритмы LRA\*[5], WHCA\*[6], FAR[7]. Алгоритмы, опирающиеся на децентрализованный подход, являются более простыми и вычислительно эффективными, однако они не являются полными. Для устранения этого недостатка был разработан алгоритм MAPP[8].

# 2. Алгоритм MAPP

Алгоритм MAPP использует децентрализованный подход к решению задачи поиска путей для множества агентов и при этом является полным для определенного класса задач, именуемого Slideable. Задача относится к классу Slideable, если  $\forall u \in U \exists \pi(u)$  и выполнены три условия:

1.  $\forall i \in (0, k_u) l_{i-1}^u, l_i^u, l_{i+1}^u \in \pi(u)$  существует альтернативный путь  $\Omega_i^u$  между вершинами  $l_{i-1}^u$  и  $l_{i+1}^u$ , который не включает в себя вершину  $l_i^u$  – см. Рис. 2.
2. На первом шаге алгоритма у каждого агента свободна вершина  $l_1^u$  (не занята другим агентом),  $l_1^u \neq l_0^v \forall u, v \in U$ .
3. (a)  $(\forall v \in U \setminus \{u\}): t_u \notin \pi(v)$  (b)  $(\forall v \in U, \forall i \in (1, \dots, k_{v-1}): t_u \notin \Omega_i^v$ ,

то есть ни один путь не может включать в себя целевые вершины других агентов.



**Рис. 2.** Альтернативный путь  $\Omega_i^u$

Алгоритм МАРР работает в два этапа:

1. Для каждого агента ищется путь с помощью модифицированного алгоритма  $A^*$ . Модификация заключается в следующем: раскрывая вершину  $x'$ , в список *open* добавляется только та смежная вершина  $x''$ , для которой существует альтернативный путь  $\Omega_x^u$  между вершинами  $x''$  и  $x$  – родительской вершиной  $x'$ . Также в список *open* не добавляются вершины, которые являются целевыми для других агентов. Подробнее о принципах работы  $A^*$  см. [<http://theory.stanford.edu/~amitp/GameProgramming/AStarComparison.html>].

2. Агенты, для которых найдены пути, осуществляют движение по ним в соответствии с заданными приоритетами. Если возникает конфликт, агент с более низким приоритетом сдвигается в сторону и ждёт пока агент с более высоким приоритетом освободит вершину, в которой возник конфликт. Благодаря зафиксированным приоритетам агенты не застревают и на каждом шаге, по крайней мере, один агент, приближается к своей цели.

### 3. Экспериментальные исследования

Алгоритм МАРР был протестирован на коллекции городских карт (100 шт.). Каждая карта соответствует области размером  $\sim 1,81 \text{ км}^2$ . Размер МТ-графа –  $501 \times 501$ . На каждой

карте определенным образом зафиксированы начальные и целевые положения для 100 агентов (см. Рис. 3).



**Рис. 3.** Городская карта в виде МТ-графа

Было проведено 2 серии экспериментов. В первом эксперименте все агенты расположены случайным образом на краях карты. При этом цель для каждого агента находится на противоположной стороне карты. Во втором эксперименте стартовая и целевая вершины для всех агентов находятся в областях размером 50x50, расположенных на противоположных сторонах карты. Для каждого графа было сгенерировано 4 задания (по 2 на эксперимент), таким образом, в общей сложности было протестировано 400 различных заданий.

В таблице 1 представлены результаты тестирования, усредненные по всем заданиям.

**Таблица 1.** Результаты тестирования

Эксперимент	Время	Память	Длина пути	Кол-во конфликтов	Решено
1	32.606	84955	540.457	33.015	100%
2	31.398	59362	540.559	192.31	99.97%

В первом эксперименте решено 100% заданий, то есть

для всех заданий построены неконфликтные траектории для всех 100 агентов. Во втором эксперименте этот показатель составляет 99,97%. На 6 заданиях алгоритм не смог построить траектории для 1 агента, которые удовлетворяли бы ограничениям класса Slideable. Вторая серия экспериментов является более сложной с точки зрения разрешения конфликтов. Все агенты движутся из одной общей области в другую и поэтому их траектории зачастую пересекаются. В связи с этим во втором эксперименте возникает почти в 6 раз больше конфликтов. Однако это почти никак не сказывается на времени, необходимом для решения задания, так как большая часть времени тратится на первый этап, когда строятся независимые траектории.

Показатель затрат памяти во втором эксперименте существенно ниже, чем в первом эксперименте. Затраты памяти равны максимальному количеству вершин, которое алгоритм хранит в памяти в процессе независимого планирования траектории для отдельного агента. В первом эксперименте все агенты располагались случайным образом, из-за чего почти во всех заданиях, по крайней мере, для одного из агентов приходилось строить траекторию, огибающую большое количество препятствий, и как следствие хранить в памяти большое число вершин. Во втором же эксперименте траектории для всех агентов были примерно одинаковыми на каждом задании. Поэтому в некоторых заданиях, где мало препятствий между стартовой и целевой областями, максимальные затраты по памяти небольшие и при усреднении повлияли на итоговые значения.

## **Выводы**

В работе рассмотрена задача планирования траекторий для коалиции агентов, проведен анализ существующих методов решения, описан и программно реализован алгоритм MAPR. Результаты экспериментов подтвердили его применимость для задач навигации коалиции агентов в городской среде.

## Список литературы

1. *J.E. Hopcroft, J.T. Schwartz, and M. Sharir.* On the complexity of motion planning for multiple independent objects; pspace-hardness of the «warehouseman's problem» // The International Journal of Robotics Research, December 1984, 3(4). P. 76–88.
2. *Яковлев К.С.* Графы специальной структуры в задачах планирования траектории // Труды III международной конференции «Системный анализ и информационные технологии САИТ-2009». М: ИСА РАН, 2009.
3. *LaValle, S., and Hutchinson, S. A.* Optimal motion planning for multiple robots having independent goals // IEEE Transactions on Robotics and Automation, 1998.
4. *Trevor Standley.* Finding optimal solutions to cooperative pathfinding problems // AAAI, 2010. P. 173–178.
5. *B. Stout.* Smart Moves: Intelligent Pathfinding // Game Developer Magazine, October/November 1996.
6. *D. Silver.* Cooperative pathfinding // AI Programming Wisdom, 2006.
7. *K.-H. C. Wang and A. Botea.* Fast and Memory-Efficient Multi-Agent Pathfinding // ICAPS, 2008. P. 380–387.
8. *Wang, K.-H. C., & Botea, A.* Tractable Multi-Agent Path Planning on Grid Maps // In Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), 2009. P. 1870–1875.

# Математическая модель для анализа характеристик интерференции при взаимодействии устройств в беспроводной сети\*

В. О. Бегишев<sup>6</sup>

**Аннотация:** Оценка отношения сигнала к интерференции плюс шум (SINR, Signal to Interference plus Noise Ratio) между взаимодействующими устройствами является одной из основных задач анализа показателей качества функционирования беспроводной сети. В работе представлена математическая модель для оценки SINR в случае, когда одно принимающее устройство (приемник) и одно передающее устройство (передатчик), образующие кластер, расположены в помещении прямоугольной формы. Такой кластер создается, например, при обмене информацией между устройствами в торговом центре, где магазины имеют регулярное расположение. В работе рассмотрен сценарий взаимодействия устройств по восходящей (uplink) линии связи. Представлены результаты численного анализа для выбранного сценария.

**Ключевые слова:** SINR; математическая модель; интерференция

## Введение

В современных беспроводных сетях, построенных на базе технологии LTE (Long Term Evolution) под интерференцией понимается взаимодействие сигналов, передаваемых разными источниками на одном и том же радиоканале. Интерференция является основным ограничивающим фактором производительности систем беспроводной радиосвязи [1]. Источниками помех могут быть мобильные устройства, находящиеся в

---

\*При финансовой поддержке РФФИ, проекты № 14-07-00090, 15-07-03051

<sup>6</sup> РУДН; 1988 г.п.; begishevu@mail.ru

той же самой соте, также совершение вызова в соседнюю ячейку или базовые станции, работающие в той же полосе частот. Явление интерференции наиболее часто наблюдается в густонаселенных районах с высокими зданиями и с большим количеством базовых станций и мобильных телефонов. Как правило, основным генератором интерференции являются передатчики от конкурирующих операторов сотовой связи, так как их базовые станции часто размещаются в непосредственной близости друг от друга для обеспечения максимального покрытия своих клиентов.

В работе исследуется SINR под упрощающим предположением о нулевой мощности шума, то есть отношение сигнала к интерференции (SIR, Signal to Interference Ratio). Среда распространения сигналов происходит в прямоугольных кластерах определенных размеров. Позиции активных мобильных пользователей являются равномерно распределены в этих кластерах [2]. Для построения математической модели применяются методы из стохастической геометрии для оценки SIR в каналах восходящей и нисходящей линий связи. Результаты исследований позволяют получить интегральные выражения для значений SIR и могут быть применены для анализа качества работы беспроводных сетей.

## 1. Описание модели

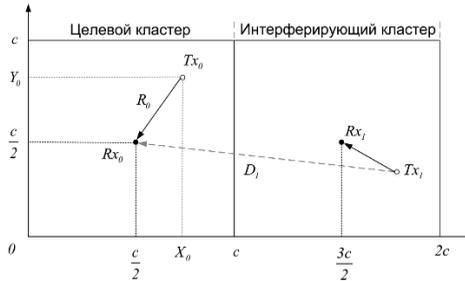
Рассмотрим случай, когда устройства абонента распложены в сетке с прямоугольными кластерами. В качестве примера можно привести торговый центр с его регулярно расположенными магазинами. В задаче рассматривается один целевой кластер и соответствующая ей пара узлов  $\langle Tx_0, Rx_0 \rangle$ , а пары устройств  $\langle Tx_i, Rx_i \rangle$ ,  $i = \overline{1,6}$  в остальных кластерах будут интерферировать с рассматриваемой парой в целевом кластере. Здесь приняты обозначения:  $Tx_i$  – передатчик и  $Rx_i$  – приемник. Общая схема взаимодействия беспроводных устройств описана в работе [3].

Так как мы предполагаем, что мощности сигналов на передатчиках одинаковые, то формула для вычисления отношения сигнала к интерференции принимает вид [4]:

$$SIR = \frac{gR_0^{-\alpha_1}}{g \sum_{i=1}^N D_i^{-\alpha_2}} = R_0^{-\alpha_1} \cdot \sum_{i=1}^N D_i^{\alpha_2}, \quad (1)$$

где  $R_0$  – расстояние между передатчиком  $Tx_0$  и приемником  $Rx_0$ ,  $D_i$  – расстояние между  $Tx_i$  и  $Rx_0$ .

Для краткости изложения рассмотрим простейший сценарий расположения устройств в кластерах для случая восходящей линии связи. Канал нисходящей линии связи исследуется аналогичным образом. Ниже приведены аналитические формулы для квадратных кластеров, представленные также в [5].



**Рис. 1.** Сценарий восходящей линии связи

Случай восходящей линии связи (uplink) представлен на Рис. 1, где рассматриваются два квадратных кластера со стороной, равной  $c$ . Один из кластеров является целевым, другой – интерферирующим. Оба приемника  $Rx_0$  и  $Rx_1$  расположены в центрах, а их передатчики равномерно распределены относительно своих приемников. В этой конфигурации передатчик  $Tx_1$  создает интерференцию приемнику  $Rx_0$ . Декартова система координат выбрана таким образом, чтобы приемник  $Rx_0$  имел координаты  $(c/2; c/2)$ . Для рассматриваемого случая,  $SIR$  принимает следующий вид:

$$SIR = R_0^{-\alpha_1} \cdot D_1^{\alpha_2}, \quad (2)$$

Отметим, что расстояния  $R_0$  и  $D_1$  являются независимыми случайными величинами, а координаты передатчиков  $Tx_0$  и  $Tx_1$  равномерно распределены в своих кластерах. Необходимо получить плотности случайной величины SIR. Задача решена с помощью метода, изложенного в [6]. Ниже представлена в интегральном виде аналитическая формула для расчета SIR, где  $x, y_0, y_1$  — являются вспомогательными переменными.

$$SIR(x) = \int_0^\infty \sum_{i=1}^6 1_{\Omega_i}(x, y_2) W_i(x, y_2) dy_2 \quad (3)$$

где  $W_i, i = \overline{1,6}$  — функции плотности SIR.

$$\begin{aligned} W_1(y_1, y_2) &= \frac{8y_2^{\frac{2}{\alpha_1}}}{\alpha_1 \alpha_2 c^4} (y_1 y_2)^{\frac{2}{\alpha_2} - 1} \arcsin \left[ \frac{\sqrt{-c^2 + 4(y_1 y_2)^{\frac{2}{\alpha_2}}}}{2(y_1 y_2)^{\frac{1}{\alpha_2}}} \right] \times \\ &\times \left( \arcsin \left[ \frac{c}{2y_2^{\frac{1}{\alpha_1}}} \right] - \arcsin \left[ \frac{\sqrt{-c^2 + 4y_2^{2/\alpha_1}}}{2y_2^{\frac{1}{\alpha_1}}} \right] \right), (y_1, y_2) \in \Omega_1 \\ W_2(y_1, y_2) &= \frac{4\pi y_2^{\frac{2}{\alpha_1}}}{\alpha_1 \alpha_2 c^4} \arcsin \left[ \frac{\sqrt{-c^2 + 4(y_1 y_2)^{\frac{2}{\alpha_2}}}}{2(y_1 y_2)^{\frac{1}{\alpha_2}}} \right], (y_1, y_2) \in \Omega_2 \\ W_3(y_1, y_2) &= \frac{8y_2^{\frac{2}{\alpha_1}}}{\alpha_1 \alpha_2 c^4} \arcsin \left[ \frac{c}{2(y_1 y_2)^{\frac{1}{\alpha_2}}} \right] \times \\ &\times \left( \arcsin \left[ \frac{c}{2y_2^{\frac{1}{\alpha_1}}} \right] - \arcsin \left[ \frac{\sqrt{-c^2 + 4y_2^{2/\alpha_1}}}{2y_2^{\frac{1}{\alpha_1}}} \right] \right), (y_1, y_2) \in \Omega_3 \\ W_4(y_1, y_2) &= \frac{4\pi y_2^{\frac{2}{\alpha_1}}}{\alpha_1 \alpha_2 c^4} \arcsin \left[ \frac{c}{2(y_1 y_2)^{\frac{1}{\alpha_2}}} \right], (y_1, y_2) \in \Omega_4 \end{aligned} \quad (4)$$

$$\begin{aligned}
W_5(y_1, y_2) &= \frac{8y_2^{\frac{2}{\alpha_1}}}{\alpha_1 \alpha_2 c^4} \left( \arcsin \left[ \frac{c}{2(y_1 y_2)^{\frac{1}{\alpha_2}}} \right] \right. \\
&\quad \left. - \arcsin \left[ \frac{\sqrt{-9c^2 + 4(y_1 y_2)^{\frac{2}{\alpha_2}}}}{2(y_1 y_2)^{\frac{1}{\alpha_2}}} \right] \right) \times \\
&\quad \times \left( \arcsin \left[ \frac{c}{2y_2^{\frac{1}{\alpha_1}}} \right] - \arcsin \left[ \frac{\sqrt{-c^2 + 4y_2^{2/\alpha_1}}}{2y_2^{\frac{1}{\alpha_1}}} \right] \right), (y_1, y_2) \in \Omega_5 \\
W_6(y_1, y_2) &= \frac{4\pi y_2^{\frac{2}{\alpha_1}}}{\alpha_1 \alpha_2 c^4} \left( \arcsin \left[ \frac{c}{2(y_1 y_2)^{\frac{1}{\alpha_2}}} \right] \right. \\
&\quad \left. - \arcsin \left[ \frac{\sqrt{-9c^2 + 4(y_1 y_2)^{\frac{2}{\alpha_2}}}}{2(y_1 y_2)^{\frac{1}{\alpha_2}}} \right] \right), (y_1, y_2) \in \Omega_6
\end{aligned}$$

где  $\Omega = \bigcup_1^6 \Omega_i$  – области интегрирования в формуле (6). В следующем разделе приведены результаты численного анализа с использованием полученных здесь формул.

## 2. Численный анализ

В этом разделе представлены результаты численного анализа рассматриваемой в работе математической модели. Корректность выведенных аналитических формул подтверждена разработанным симулятором, описанным в [3,7]. Графически показан характер поведения интерференции в зависимости от входных параметров. Исходные данные представлены в описании графиков – это размеры кластеров и значения экспонент затухания  $\alpha_1$  и  $\alpha_2$ . На Рис. 2 при размерах кластера  $c = 10$  исследованы характеристики случайной величины SIR в зависимости от различных значений  $\alpha_2$ . Приходим к выводу, что чем больше  $\alpha_2$ , тем лучше становится интерференционная картина на стороне целевого приемника. Заметим, что в модели

значение  $\alpha_2$  определяет материал стен. Рассматриваемый сценарий взаимодействия устройств имеет ряд интересных свойств, в частности, значения SIR не зависят от размера кластера, если он имеет квадратную конфигурацию. Таким образом, наша модель может быть полезна для анализа взаимодействия беспроводных устройств.

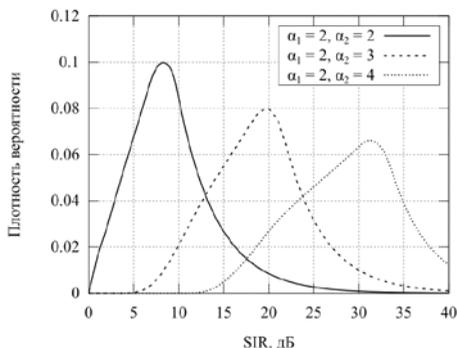


Рис. 2. Зависимость от экспоненты затухания при  $c = 10$

## Выводы

В этой работе построена математическая модель по описанию характера поведения SIR в двух соседних квадратных кластерах. Установлено, что для восходящей линии связи распределение отношения сигнал к интерференции можно оценить аналитически без необходимости затрат времени на имитационное моделирование. Численное исследование показывает, что для квадратных конфигураций кластеров значение SIR является нечувствительным к длине стороны квадрата. Кроме того, экспонента потерь  $\alpha_2$  существенно зависит от материала стенки смежных комнат. Наконец, в модели получено распределение SIR только для одного приемника, который создает интерференцию. В реальных сценариях таких устройств, образующих помеху, достаточно много. Поэтому разработка аналитической модели для этого случая является предметом наших текущих исследований.

## Список литературы

1. *Самуйлов А. К., Гайдамака Ю. В.* Метод расчета характеристик интерференции двух взаимодействующих устройств в беспроводной гетерогенной сети // Информатика и ее применения, 2015. Т.9. No.1. С. 9-14.
2. *Самуйлов А. К., Бегишев В.О., Ковальчуков Р.Н.* Расчет интерференции при взаимодействии беспроводных устройств в квадратных кластерах // Труды всероссийской конференции «Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем». –М.: РУДН, 2015, С. 110.
3. *Begishev V., Kovalchukov R., Samuylov A., Ometov A., Moltchanov D., Gaidamaka Yu., Andreev S., and Koucheryavy Y.* On Numerical Estimation of SINR for Square Wireless Clusters // Proceedings of the IX International Workshop APTP+MS' 2015, August 10 – 13, 2015, Tampere, Finland. – Finland, Tampere, Printed at TUT. – 2015. – P.45-47.
4. Rappaport, T.S Wireless communications: principles and practice. et al.: Prentice Hall PTR, New Jersey, 1996. – vol. 2 – 640 p.
5. *Begishev V., Kovalchukov R., Samuylov A., Ometov A., Moltchanov D., Gaidamaka Yu., Andreev S., Koucheryavy E.* An analytical approach for SINR estimation in adjacent rectangular premises // Proc. of the 15th IEEE International Conference on Next Generation Wired/Wireless Networking NEW2AN 2015, – 2015. pp.446-458.
6. *Левин Б. П.* Теоретические основы статистической радиотехники. – М.: Радио и связь, 1989. – 656 с.
7. *Гайдамака Ю. В., Самуйлов А. К., Бегишев В. О., Ковальчуков Р.Н.* Имитационная модель для оценки интерференции при взаимодействии беспроводных устройств в прямоугольных кластерах // Технологии информационного общества. – М.: Т-Comm – Телекоммуникация и транспорт, 2015. С. 7.

# Оценка качества аппроксимации хардкор процессов Матерна пространственными пуассоновскими процессами\*

В. А. Бесчастный<sup>7</sup>

**Аннотация:** Наиболее точные результаты при моделировании беспроводных сетей позволяют достичь хардкор процессы Матерна. Тем не менее они достаточно сложны для анализа и потому часто вместо них применяют пуассоновские пространственные процессы, которые отличаются своей простотой и широким применением в различных направлениях исследований. При этом переходе происходит некоторая потеря точности. В работе проводится оценка целесообразности такого перехода в контексте беспроводных технологий передачи данных, в которых ключевым показателем качества функционирования является отношение сигнала к интерференции SIR между взаимодействующими устройствами. В работе представлена математическая модель фрагмента беспроводной сети для оценки замены процессов. Для модели представлены результаты сравнительного численного анализа.

**Ключевые слова:** пространственные процессы; SIR; D2D

## Введение

Установление прямых соединений между близко расположенными мобильными устройствами позволяет более эффективно использовать выделенный спектр частот, увеличить общую пропускную способность сети, а также позволяет предоставлять услуги на базе технологии P2P (Peer-to-Peer) и определения текущего местоположения (Location-based services). Еще одним преимуществом такого типа соединений является

---

\*Работа выполнена при поддержке РФФИ в рамках научных проектов № 14-07-00090, 15-07-03608, 15-07-03051.

<sup>7</sup> РУДН; 1992 г.р.; vbeschastny@sci.pfu.edu.ru

возможность организации с помощью LTE-устройств (Long Term Evolution), поддерживающих технологию D2D (Device-to-Device), сетей общественной безопасности, которые должны функционировать там, где сотовая связь по тем или иным причинам недоступна, например, в случае возникновения аварийных ситуаций. Внедрение D2D-технологий постепенно ставит вопрос о пересмотре традиционной архитектуры сотовых сетей, центральными элементами которых являются базовые станции [1].

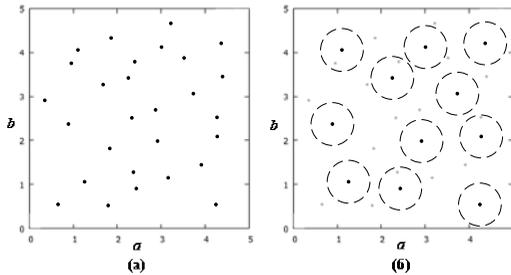
При организации D2D-соединений возникает множество требующих решения задач. Одной из наиболее принципиальных задач является разделение частотного ресурса между двумя типами соединений: D2D-соединениями и соединениями традиционной сотовой связи. Совокупность различных аспектов, в том числе интерференция и безопасность, не позволяет с уверенностью дать ответ: должны ли D2D-устройства использовать строго выделенный для них спектр частот, или у них может быть доступ к частотам, используемым устройствами сотовой связи. К тому же, задача распределения частот осложняется тем, что устройства в некоторых случаях могут переключаться между обоими типами соединений.

Как правило, для моделирования беспроводных сетей используются пространственные процессы (Point Processes). Пространственный процесс является случайным процессом и представляет собой множество точек, распределенных в некотором  $n$ -мерном пространстве. Одно из наилучших приближений к реальным данным позволяет получить хардкор процессы Матерна (МНС, Matern Hardcore Process), в которых устанавливается некоторое минимально возможное расстояние между произвольной парой устройств. Причем в большинстве случаев МНС II-типа дает лучшее приближение, чем МНС I-типа [2]. Однако такие процессы очень сложны для анализа, поскольку для них не существует достаточного количества аналитических выражений. Часто их заменяют хорошо изученными и относительно простыми пространственными пуассоновскими процессами (PPP, Poisson Point Process) [3].

В работе исследуется возможность и точность аппроксимации хардкор процессов Матерна пространственными пуассоновскими процессами. Оценка точности построена на анализе расстояний между устройствами, а также отношения сигнала к интерференции (SIR, Signal to Interference) [4]. В разделе 1 рассмотрена типовая модель фрагмента сети с установлением прямых соединений. Раздел 2 содержит результаты численного анализа. Основные сравнительные характеристики пространственных процессов рассмотрены в разделе 3.

## 1. Описание модели

Рассмотрим фрагмент сети, который представляет собой плоскость с распределенными на ней устройствами (Рис. 1). Для простоты будем считать, что на плоскости нет объектов, которые могли вызвать затухания сигнала, мощность сигнала всех устройств положим равной. В рамках данной работы рассматривается только нисходящий сигнал от устройств-источников, которые в процессе функционирования могут вызывать явление интерференции. Таким образом, имеет смысл рассматривать все устройства, как потенциальные источники.



**Рис. 1.** Фрагмент сети с беспроводными устройствами, смоделированный по PPP (а) и МНС (б)

Из представленных упрощений можно получить формулу для определения SIR в виде (1).

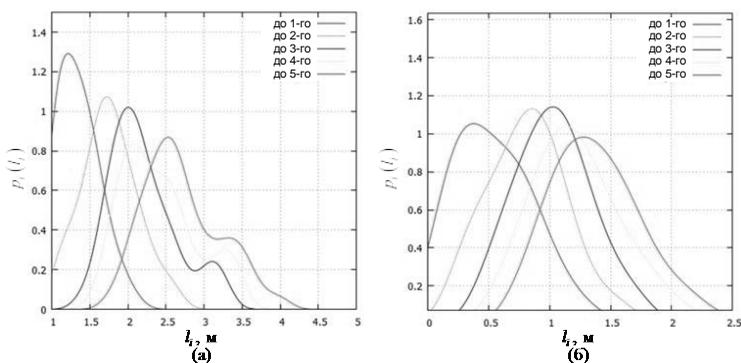
$$SIR = \frac{g l_0^{-\alpha}}{g \sum_{i=1}^N l_i^{-\alpha}} = l_0^{-\alpha} \cdot \left[ \sum_{i=1}^N l_i^{-\alpha} \right]^{-1}, \quad (1)$$

где  $\alpha$  – показатель затухания сигнала в среде (далее в работе

полагается  $\alpha = 2$ , что соответствует наиболее идеальным условиям),  $l_0$  - расстояние между рассматриваемой парой целевой источник-приемник (ЦИ-П),  $l_i$  - расстояние между парами целевой источник-интерферирующий  $i$ -источник (ЦИ-ИИ),  $g$  - мощность сигнала от источников,  $N + 1$  - общее количество источников.

## 2. Численный анализ

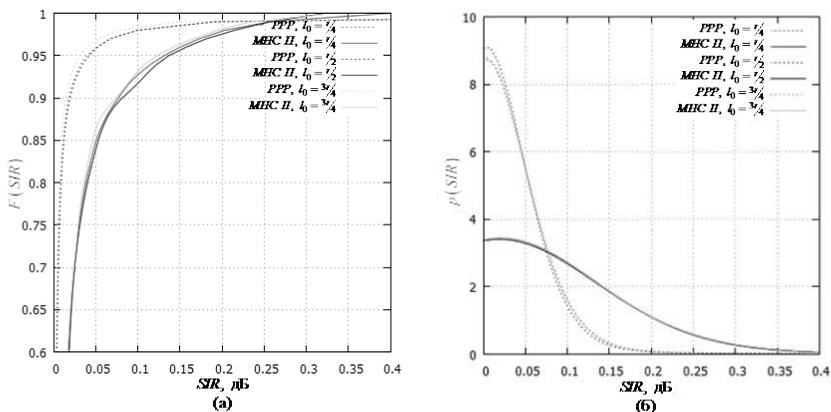
Как уже ранее отмечалось, наиболее интересным параметром при исследовании взаимодействия беспроводных устройств является SIR. Из формулы (1) видно, что наибольший вклад в уменьшение этого показателя привносят ближайшие устройства, поэтому целесообразно проводить анализ в контексте именно таких устройств. Для определенности будем рассматривать пять пар ЦИ-ИИ с наименьшими расстояниями. На Рис. 2 представлены эмпирические функции плотности распределения расстояний, из которых видно, что отличия в расстояниях, даже в среднем значении, существенны.



**Рис. 2.** Эмпирические функции плотности распределения расстояний до ближайших пяти ИИ при (а) МНС II, (б) PPP

Результаты показаны для пространственных процессов с интенсивностями  $\lambda_{MNC} = \lambda_{PPP} = 1$  ИИ/м<sup>2</sup>, параметром МНС-процесса  $r = 1$ .

На Рис. 3 представлена эмпирические функции распределения и функция плотности SIR для пары ЦИ-П, где источник находится в трех положениях:  $l_0 \in \left\{ \frac{1}{4}r; \frac{1}{2}r; \frac{3}{4}r \right\}$ . Параметры моделирования аналогичны.



**Рис. 3.** Эмпирические функция распределения (а) и функция плотности (б) SIR для ЦИ-П

Таким образом, можно сделать вывод о том, что в терминах SIR погрешность при переходе между процессами усугубляется еще больше.

В качестве критерия целесообразности перехода от МНС к PPP выберем критерий Фишера равенства выборочных дисперсий SIR [5]. В случае если критерий Фишера выполняется при некотором наборе исходных параметров, то переход между процессами не приведет к значительной потери точности. В случае, если критерий не выполняется, переход приведет к резкому ухудшению приближенности моделируемых процессов к реальным. На Рис. 4 представлена диаграмма, на которой область в левом углу соответствует параметрам, на которых критерий выполняется полностью. Для двух зон в центре критерий не выполняется только для не более чем двух и четырёх ближайших пар ЦИ-ИИ соответственно.

Стоит отметить, что по оси абсцисс отложены значения

параметра  $\gamma = \frac{r}{S}$ , где  $S$  — это площадь исследуемого фрагмента сети. Несложно показать, что при выполнении этого соотношения, результаты являются действительными для любых плоскостей выпуклой формы, подобной квадрату. Исключением являются слишком малые плоскости, площадь которых не позволяет разместить статистически достаточное количество устройств (приблизительно менее  $25 \text{ м}^2$ ).

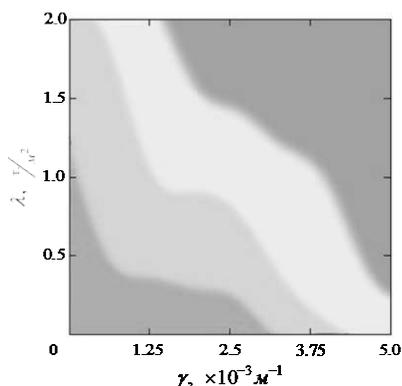


Рис. 4. Критерий Фишера равенства выборочных дисперсий SIR

## Выводы

В работе построена математическая модель взаимодействия беспроводных устройств. Установлено, что аппроксимация хардкор процессов Матерна пространственными пуассоновскими процессами приводит к значительной потере точности при моделировании процессов. Тем не менее, существует ограниченная область значений параметров системы, при которой такой переход не влечет за собой качественного ухудшения приближенности модели к реальным процессам, и при этих значениях пуассоновские пространственные процессы могут применяться с большой долей уверенности.

## Список литературы

1. *Hesham ElSawy, Ekram Hossain, Mohamed-Slim Alouini* Analytical Modeling of Mode Selection and Power Control for Underlay D2D Communication in Cellular Networks // IEEE Ieee Transactions on Communications, Vol. 62, No. 11. – 2014, pp.4147–4161.
2. *Qiaoyang Ye, Mazin Al-Shalash, Constantine Caramanis, Jeffrey G. Andrews* Device-to-Device Modeling and Analysis with a Modified Matern Hardcore BS Location Model // IEEE Communication Theory Symposium. – 2013, pp.1825-1830.
3. *F. Baccelli, B. Blaszczyszyn* Stochastic geometry and wireless networks, vol. 1. -- Now Publishers Inc, 2009. – 150 с.
4. *Begishev V., Kovalchukov R., Samuylov A., Ometov A., Moltchanov D., Gaidamaka Yu., Andreev S., Koucheryavy E.* An analytical approach for SINR estimation in adjacent rectangular premises // Proc. of the 15th IEEE International Conference on Next Generation Wired/Wireless Networking NEW2AN 2015. – 2015, pp.446–458.
5. *Левин Б. П.* Теоретические основы статистической радиотехники. – М.: Радио и связь, 1989. – 656 с.

# Метод одновременного картирования и локализации по видеопотоку, полученному с единственной камеры, LBA-vSLAM\*

А.В. Боковой<sup>8</sup>

**Аннотация:** В работе представлен новый метод одновременного картирования и локализации по видеопотоку, полученному с единственной камеры. Задача картирования и локализации рассмотрена в контексте более общей проблемы автономной навигации беспилотного летательных аппарата (БПЛА).

**Ключевые слова:** компьютерное зрение; методы одновременного картирования и локализации; видеопоток.

## Введение

В настоящее время многие робототехнические системы наряду с инерциальными датчиками, инфракрасными и лазерными дальномерами, также оснащаются видео-датчиками (камерами). Более того всё больше и больше систем, из-за ограничений на грузоподъемность и энергоемкость, оснащаются лишь камерами. Таким образом, весьма актуальной задачей современной интеллектуальной робототехники является задача одновременного картирования и локализации по видеопотоку. На текущий момент уже существуют методы, позволяющие в той или иной степени решить эту задачу. Так, методы, основанные на поиске и сопоставлении ключевых точек, являются быстрыми, но не точными, а построенная карта обладает низкой детализацией [1, 2]. Методы, использующие всю информацию изображения, позволяют строить высокодетализированные

---

\*Работа выполнена при поддержке проекта РФФИ №15-37-51098\_МОЛ\_НР

<sup>8</sup> ФИЦ ИУ РАН (ИСА РАН); 1993 г.р.; bokovoy@inbox.ru

карты, но не работают в реальном времени без ускорения на GPU [3].

В рамках данной работы решается задача построения карты неизвестной местности и пройденной траектории робототехнической системы по ряду изображений (видеопотоку), полученному с единственной камеры. Ожидается, что метод позволит строить карты высокой детализации в реальном времени.

Итак, пусть видеопоток задан в виде:

$$\dot{I} = (I_1, I_2, \dots, I_T) \quad (1)$$

где  $I_t$  – матрица, размером  $K \times L$ , элементам которой соответствуют пиксели изображения, полученного в момент времени  $t$ , в градации серого.  $K, L, T \in \mathbb{N}$ ,  $t \leq T$ .

Для каждого  $I_t$  необходимо вычислить матрицу внешней калибровки камеры:

$$E_t = \begin{pmatrix} R & T \\ 0 & 1 \end{pmatrix} \quad (2)$$

где  $R$  – матрица поворота размером  $3 \times 3$ ,  $T$  – матрица переноса размером  $3 \times 1$ .

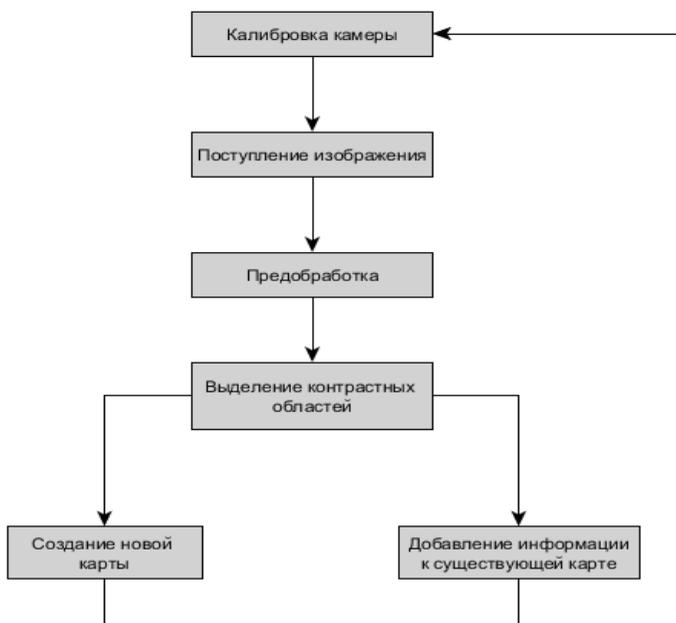
Набор матриц  $E = (E_1, \dots, E_T)$  представляет собой позиции камеры в глобальной системе координат в каждый момент времени  $t$ , что соответствует траектории, пройденной камерой за промежутки времени  $(1, T)$ .

Также, для заданного видеопотока (1), необходимо вычислить вектор  $\dot{M} = (M_1, \dots, M_P)$ ,  $P \in \mathbb{N}$ , где  $M_i$  – трехмерная координата пикселя из видеопотока (1). Полученный набор координат является трехмерной реконструкцией проекции местности, представленной на изображениях  $I_t$ .

## 1. Метод LBA-vSLAM

Предлагаемый алгоритм состоит из 5-ти основных шагов (см. Рис. 1):

1. Калибровка камеры (получение матрицы внутренней калибровки).
2. Получение изображения.
3. Предобработка изображения, устранение шумов.
4. Выделение областей изображения, по которым будет составлена карта и по которым будет производиться поиск пройденной камерой траектории.
5. Создание новой карты, при условии, что не удастся найти соответствие выделенной области на предыдущих изображениях или добавление новых элементов карты в противном случае.



**Рис. 1.** Блок-схема алгоритма LBS-vSLAM

## 1.1. Калибровка камеры

Для восстановления 3D-структуры местности по

двум или более изображениям, необходимо знать внутренние параметры камеры (фокусное расстояние, углы наклона пикселей и т.д.). Как правило, эти параметры вычисляются с помощью алгоритмов калибровки, использующих заранее известные изображения (например, изображений типа шахматная доска [[http://docs.opencv.org/doc/tutorials/calib3d/camera\\_calibration/camera\\_calibration.html](http://docs.opencv.org/doc/tutorials/calib3d/camera_calibration/camera_calibration.html)]).

В итоге, строится матрица внутренней калибровки камеры

$$A = \begin{bmatrix} \alpha_x & \gamma & u_0 \\ 0 & \alpha_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (3)$$

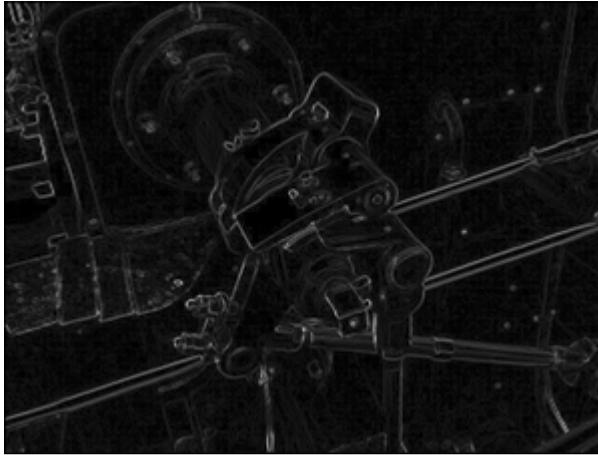
Здесь  $\alpha_x$ ,  $\alpha_y$  - фокусные расстояния по оси  $x$  и  $y$  соответственно,  $u_0$ ,  $v_0$  - координаты точки пересечения плоскости изображения с оптической осью,  $\gamma = \alpha_y \tan \psi$ ,  $\psi$  - угол наклона пикселя.

Параметры внутренней калибровки камеры вычисляются один раз и могут быть использованы повторно для одной и той же камеры.

## 1.2. Выделение областей

Для оптимизации времени работы алгоритма, предлагается использовать не всю информацию на изображении, а только контрастные области. Такой подход позволит уменьшить вычислительные затраты на исполнения алгоритма, т.к. поиск соответствий 2-х изображений будет происходить только по точкам, принадлежащим этим областям.

Для получения подобных областей предлагается использовать оператор Собеля [4]. На Рис. 2 белыми пикселями обозначены выделенные области.



**Рис. 2.** Области, выделенные с помощью фильтра Собеля.

Таким образом, для каждого изображения  $I_i$  из видеопотока  $\dot{I}$ , получаем соответствующие контрастные области  $S_i$ .

### **1.3. Создание и обработка карты с помощью метода Local Bundle Adjustment**

В начальный момент времени, когда карта отсутствует или при условии того, что на текущем изображении нет соответствий предыдущим изображениям, создается новая карта. В таком случае, первый кадр  $I_j$ , удовлетворяющий вышеописанному условию, принимается за стартовый и начинается процесс инициализации карты. Для  $S_i$  задается случайная карта глубины  $D_j$ . Далее, для областей  $S_j, S_{j+1}, \dots, S_{j+h}$ ,  $h \in \mathbb{N}$  решается задача поиска соответствующих пикселей. Предположим, что найденные соответствия  $x_j, x_{j+1}, \dots, x_{j+h}$  являются проекцией трехмерной точки  $X_p$  на изображения  $I_j, I_{j+1}, \dots, I_{j+h}$ .

Таким образом, задача сводится к минимизации проективной ошибки:

$$\Delta x_j(A, E_j, S_j X_p) = x_j - x(A, E_j, S_j X_p), \quad (4)$$

$$\Delta x_j(A, E_j, S_j X_p) \rightarrow \min, \quad (5)$$

где  $\Delta x_j$  – ошибка проекции точки  $X_p$  на изображение  $I_j$ ,  $x(A, E_j, S_j X_p)$  – предсказанное положение точки  $X_p$  на изображении  $I_j$ . Для решения этой задачи подходит метод наименьших квадратов. В рамках работы был выбран взвешенный метод Ньютона-Гаусса [5].

В случае, когда карта уже имеется, и на обработку поступает следующий кадр, минимизация проективной ошибки происходит только для соседнего кадра  $I_{j-1}$ .

## 2. Коррекция карты и пройденной траектории

При составлении карты и построении пройденной траектории, накапливается погрешность. Чем больше изображений обработано, тем больше эта ошибка. Для коррекции планируется использовать алгоритм оптимизации Pose-Graph Loop-Closure. Метод позволяет производить коррекцию и распознавать замыкания траектории, что позволит повысить точность работы алгоритма LBA-vSLAM.

Также, при распознавании замыкания и при повторном прохождении траектории, не будет необходимости заново производить инициализацию карты. Локализацию камеры можно производить на уже известной карте, что ускорит работу алгоритма при однократном или многократном прохождении одной и той же траектории.

Корректировка карты может производиться как во время основных вычислений, так и в фоновом режиме. Первый подход может применяться, если корректировка карты и пройденной траектории нужна в каждый момент

времени, например, при сильной зашумленности изображений или недостатке информации на изображении. Второй подход позволит распараллелить работу алгоритма, тем самым сократив нагрузку на вычислитель, что может быть критично при исполнении алгоритма на бортовом компьютере робототехнического устройства.

## Выводы

В статье представлен алгоритм одновременного картирования и локализации по видеопотоку LBA-vSLAM. Отличительной особенностью алгоритма является использование контрастных областей изображений для составления карты и применение алгоритма Local Bundle Adjustment для нахождения пройденной траектории.

## Список литературы

1. *Davidson A., Reid I., Stasse O.* MonoSLAM: Real-Time Single Camera SLAM // IEEE Transaction on Pattern Analysis and Machine Learning archive, 2007. Т. 29, №6, С. 1052-1067
2. *Klein G., Murray D.* Parallel Tracking and Mapping for Small AR Workspaces // In Proc. International Symposium on Mixed and Augmented Reality, 2007.
3. *Glocken B., Kelly P.H.J., Davison A. J.* Dense planar SLAM // In Proc. International Symposium on Mixed and Augmented Reality, 2014. С. 157-164
4. Vincent O.R., Folorunso O. A Scriptive Algorithm for Sobel Image Edge Detection // In Proc. Of Informing Science & IT Education Conference (InSITE), 2009.
5. Ерина М., Ю., Измаилов А., Ф. Метод Гаусса-Ньютона для отыскания особых решений систем нелинейных уравнений // Журнал вычислительной математики и математической физики, 2007. Т. 47, №6, С. 784-795

# Применение метода МГУА на основе интенсивности поисковых запросов в сети Интернет для прогноза рынка недвижимости\*

А. В. Болдырева<sup>9</sup>

**Аннотация:** В статье показана возможность прогноза цен на рынке недвижимости Москвы на основе статистики поисковых запросов в поисковом сервисе Яндекс. Модели строятся с использованием метода группового учета аргументов. Эксперименты показали ошибку в 2-3% для расчета текущей стоимости недвижимости разного типа.

**Ключевые слова:** поисковые запросы; МГУА; Интернет

## Введение

В настоящее время использование динамики поисковых запросов набирает все больший оборот. Так, в 2009 году Хьюнъянг Чой и Хэл Вэриан проанализировали такие экономические индикаторы, как автомобильные продажи, безработица и потребительское доверие [1]. Среди множества последовавших исследований можно отдельно отметить анализ экономики Тобиясом Прайзом с применением лагирования [2] и прогнозирование безработицы Никосом Аскитасом и Клаусом Зиммерманом [3]. В России подобные исследования практически не проводятся.

В статье предлагается использовать поисковые запросы для прогноза средних цен на недвижимость в Москве. Методы основаны преимущественно на тренде цен и информации по экономической конъюнктуре [4].

---

<sup>9</sup> МФТИ; РАНХиГС; 1993 г.р.; [anna.boldyreva@phystech.edu](mailto:anna.boldyreva@phystech.edu)

Анализ объемов продаж и цен на недвижимость проводился ранее, к примеру, Линн Ву, Эриком Бриньолфссоном [5]. Однако для анализа в этой статье впервые применяется описанный метод и впервые исследуется рынок в России.

Для построения объясняющих переменных используется идея Михаила Столбова по формированию барометров финансовых показателей из поисковых запросов [6]. Прогноз проводится по модифицированному методу, предложенному автором ранее для построения прогностических моделей индикаторов экономической, социальной и финансовой конъюнктуры России [7–9].

## **1. Исходные данные**

### **1.1. Индикаторы**

В качестве индикаторов в текущем исследовании использовалась статистика средних цен в рублях на кв. метр, размещенная в базе ЦИАН. Индикаторы: средние цены на покупку квартир, офисов и торговых площадей; средние цены на аренду офисов, торговых площадей, однокомнатных и двухкомнатных квартир. База данных: с августа 2013 года по август 2015 года, что вызвано ограничением сервисом Яндекса.

### **1.2. Барометры**

Базы помесечных относительных объемов запросов получены из базы сервиса статистики Яндекс. Выбор запроса определялся на основе раздела «По словам» поискового сервиса статистики Яндекс «Вместе с этим словом ищут», сервиса Google Correlate.

Отбирались слова (1005 слов и словосочетаний) с высоким коэффициентом корреляции Пирсона (больше 60%) для каждого индикатора с лагами до трех месяцев. Затем данные нормировались и для каждого лага и индикатора вычислялись средние запросы – барометры (см. данные в Таблице 1). В мо-

делях применялись линейные полиномы барометров и полиномы квадратных корней барометров.

**Таблица 1.** Примеры запросов, вошедших в барометры

<b>Индикатор</b>	<b>Запросы (дескрипторы)</b>
Квартиры	циан вторичка; недвижимость москва официальный сайт вторичка; цены на аренду квартир в москве 2015;
Аренда однокомнатных квартир	узбекско русский переводчик; квар- тиры в подмосковье; квартиры в аренду в москве от хозяев
Аренда торговых пло- щадей	риелтер; взять в аренду; кредит в за- лог недвижимости

## **2. Прогноз**

### **2.1. Метод**

Для прогнозирования применялся метод группового учета аргументов (МГУА) [10], реализованный в программе GMDH Shell. Последний месяц исключался из выборки для дальнейшей проверки качества прогнозирования.

Шаги метода:

1. Определяется серия моделей возрастающей сложности.
2. Экспериментальные данные делятся на обучающую и экзаменационную выборки.
3. Для заданного уровня сложности определяется оценка параметров модели на наборе данных с использованием внутреннего критерия.
4. Модель верифицируется на втором наборе данных с помощью внешнего критерия проверки качества;
5. Поиск останавливается при достижении минимума внешнего критерия.

Использованные алгоритмы построения моделей: комбинаторный (перебор всех моделей) и нейронный (полиноми-

альная нейронная сеть, где в каждом слое сети добавляются эффективные регрессионные модели предыдущего слоя). Выставлялось ограничение на начальную ширину слоя 5 и максимальное количество слоев 6.

Для валидации модели применялась перекрестная проверка с количеством частей  $k=2$ .

В качестве внутреннего критерия был использован МНК, в качестве внешнего – RMSE:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{24} (\hat{y}_t - y_t)^2}{24}}$$

## 2.2. Результаты

Было построено 36 моделей: по 4 модели на основе 8 барометров для каждого из 9 индикаторов.

Применялось два критерия оценки построенной модели: средняя абсолютная процентная ошибка:

$$MAPE = \frac{1}{24} \sum_{t=1}^{24} \left| \frac{\hat{y}_t - y_t}{y_t} \right| * 100\%$$

и процентное отклонение прогноза на один месяц вперед от реального значения за этот месяц:

$$P = \left( \frac{\hat{y}_{25} - y_{25}}{y_{25}} \right) * 100\%.$$

**Таблица 2.** MAPE лучших моделей

Индикатор	MAPE лучшей модели	Модели
Квартиры	2,2%	Все 4
Первичный рынок	1,6%	Нейросеть без корней
Вторичный рынок	1,1%	Нейросеть с корнями
Офисы	4,0%	Нейросеть без корней
Торговые площади	7,1%	Комби с корнями
Аренда 1-к квартир	2,1%	Нейросеть без корней, комби без корней
Аренда 2-к квартир	1,2%	Нейросеть без корней
Аренда офисов	4,0%	Нейросеть без корней

Аренда торговых площадей	11.8%	Комби без корней
--------------------------	-------	------------------

Таблица 2 содержит значения MAPE лучшей модели для каждого индикатора. Лучшие значения MAPE колеблются от 1 до 12%. В целом же значения, за исключением данных для торговых площадей и офисов, составляют 1-4%. Как можно заметить, в среднем меньшие значения MAPE дают модели, построенные с помощью нейронного алгоритма, на вход которого подаются линейные переменные.

Таблица 3 содержит процентные отклонения прогноза на 1 месяц вперед от реального значения за этот месяц.

**Таблица 3.** Р лучших моделей

Индикатор	Р лучшей модели	Тип модели
Квартиры	-4,7%	Все 4
Первичный рынок	4,6%	Нейросеть с корнями
Вторичный рынок	3,1%	Нейросеть с корнями
Офисы	-5.0%	Нейросеть с корнями, комби с корнями
Торговые площади	-11.8%	Комби без корней
Аренда 1-к квартир	2,4%	Комби без корней
Аренда 2-к квартир	-0,3%	Комби без корней
Аренда офисов	7,1%	Нейросеть без корней
Аренда торговых площадей	-14,4%	Комби без корней

Лучшие значения Р колеблются от 0 до 12%.

## **Выводы**

Исследования показали, что поисковые запросы применимы для прогнозирования цен на некоммерческую недвижимость в Москве. Модели, описанные в статье, дают ошибку прогнозирования 2-3%.

## **Благодарности**

*Автор выражает благодарность Александрову Михаилу*

*Ароновичу, научному руководителю, и Кошулько Алексею Анатолиевичу, одному из авторов программы и консультанту по GMDH Shell.*

## **Список литературы**

1. *Choi H., Varian H.* Predicting the Present with Google Trends. // Working paper, Google Inc., 2009.
2. *Preis T., Reith D., Stanley H.E.* Complex dynamics of our economic life on different scales: insights from search engine query data. // Phil. Trans. R. Soc. A, 2010. Pp. 5707-5719.
3. *Askitas N., Zimmermann K.F.* Google econometrics and unemployment forecasting, // Technical report / SSRN 899, 2010.
4. *Стерник Г. М.* Методика среднесрочного прогнозирования развития рынка жилья города (региона) // Имущественные отношения в Российской Федерации, 2012. No.9.
5. *Wu L., Brynjolfsson E.* The future of prediction: How Google searches foreshadow housing prices and sales // Economics of Digitization. Chicago: University of Chicago Press, 2013.
6. *Столбов М.И.* Статистика поиска в Google как индикатор финансовой конъюнктуры // Вопросы экономики, 2011. No. 11.
7. *Болдырева А.В.* Обработка поисковых запросов в сети Интернет для анализа и прогноза социально-экономических показателей // Сборник студенческих работ РАНХиГС, 2015.
8. *Болдырева А.В.* Анализ экономической, социальной, демографической ситуации в стране по интенсивности запросов в поисковой сети Интернет // Сборник студенческих работ РАНХиГС, 2015.
9. *Болдырева А.В.* Построение прогнозных моделей экономической и социальной конъюнктуры по интенсивности запросов в поисковой сети Интернет // Сборник студенческих работ РАНХиГС, 2015.
10. *Ивахненко А.Г., Степашко В.С.* Помехоустойчивость моделирования: монография // Киев: Наукова думка, 1985. – 216 с.

# Оценка многомерной плотности распределения

А.И. Борисов<sup>10</sup>

**Аннотация:** Данная работа посвящена проблеме оценки вероятностных плотностей с использованием аппарата вейвлет-анализа. Вейвлет-анализ – сравнительно недавно разработанный аппарат. Свойства локализации вейвлет-функций во времени и пространстве делают вейвлет-оценки конкурентоспособными. В работе представлены метод оценки одномерных плотностей, а также возможный метод обобщения его на многомерный случай с помощью преобразования Радона и его обращения. Кроме того, проведена реализация метода для двумерного случая и представлены результаты его работы.

**Ключевые слова:** вейвлет-анализ; преобразование Радона; плотность распределения

## Введение

Рассматривается проблема непараметрического оценивания многомерной вероятностной плотности, актуальная, например, в задачах классификации и распознавания изображений.

Данная работа представляет собой описание метода, позволяющего построить оценку многомерной плотности. Также проведена работа по реализации данного метода.

Метод основан на использовании проекций наблюдений случайной величины на некоторые направления. С их помощью можно построить оценки одномерных плотностей. Из них с помощью методов обращения преобразования Радона можно получить исходную плотность распределения.

Метод был реализован в среде Matlab. Он был протест-

---

<sup>10</sup> ВМК МГУ им. М.В. Ломоносова; 1995 г.р.; artvenom1@mail.ru

стирован на нормальном распределении и распределении Стьюдента в двумерном случае, показав неплохие результаты.

В настоящее время исследуются свойства оценок, построенных по указанному методу. В частности, ожидаются хорошие результаты для плотностей, не обладающих локальной гладкостью.

## 1. Вейвлеты

Вейвлеты представляют собой семейство базисных функций, позволяющие аппроксимировать другие функции. Вейвлет-базисы предлагают достаточно хорошие свойства, такие, как определенная степень гладкости, локализация во времени и пространстве. Кроме того, существуют довольно быстрые алгоритмы разложения функции по коэффициентам вейвлет-базиса, а также для реконструкции функции по этим коэффициентам. Это делает вейвлеты пригодными для непараметрического оценивания плотности распределения.

Пусть имеется неоднородный ортонормированный вейвлет-базис  $\{\varphi_{j_0,k}(x) = 2^{j_0/2}\varphi(2^{j_0}x - k), k \in \mathbb{Z}\}$  и  $\{\psi_{j,k}(x) = 2^{j/2}\psi(2^jx - k), k \in \mathbb{Z}, j \geq j_0\}$ , где  $\varphi(x)$ ,  $\psi(x)$  – масштабирующая и вейвлет-функции соответственно. Предположим, что эти функции достаточное число раз дифференцируемы и имеют компактный носитель,  $\varphi(x)$  нормирована, то есть ее интеграл на всей числовой прямой равен 1.

## 2. Описание метода для одномерных плотностей распределения

### 2.1. Построение оценки

Рассмотрим независимые одинаково распределенные наблюдения  $X_1, X_2, \dots, X_N$  случайной величины, имеющей плотность распределения  $f(x)$ . Тогда  $f(x)$  имеет формальное вейвлет-разложение

$$f(x) = \sum_k c_{j_0,k} \varphi_{j_0,k}(x) + \sum_{j_0}^{+\infty} \sum_k d_{j,k} \psi_{j,k}(x),$$

где

$$c_{j_0,k} = \int_{-\infty}^{\infty} f(x) \varphi_{j_0,k}(x) dx,$$

$$d_{j,k} = \int_{-\infty}^{\infty} f(x) \psi_{j,k}(x) dx.$$

Поскольку  $f(x)$  – плотность распределения, то для нее можно записать вейвлет-оценку [1]

$$\hat{f}(x) = \sum_k \hat{c}_{j_0,k} \varphi_{j_0,k}(x) + \sum_{j_0}^{j_1} \sum_k \hat{d}_{j,k} \psi_{j,k}(x), \quad (1)$$

где

$$\hat{c}_{j_0,k} = n^{-1} \sum_{i=1}^n \varphi_{j_0,k}(X_i),$$

$$\hat{d}_{j,k} = n^{-1} \sum_{i=1}^n \psi_{j,k}(X_i),$$

$j_0, j_1$  выбраны некоторым надлежащим образом. Нетрудно видеть, что  $\hat{c}_{j_0,k}$ ,  $\hat{d}_{j,k}$  – несмещенные оценки для  $c_{j_0,k}$  и  $d_{j,k}$  соответственно.

### 3. Переход к многомерному случаю

Преобразованием Радона в  $d$ -мерном пространстве функции  $f: R \rightarrow R^d$  называется функция

$$R_\theta f[s] = \int_{x^* \theta = s} f(x) dx,$$

где  $\theta$  – произвольный единичный вектор. Формально для фиксированных  $\theta, s$  эта величина имеет смысл интеграла функции  $f$  по гиперплоскости, перпендикулярной  $\theta$  и удаленной на  $s$  от начала координат.

Рассмотрим Сведение задачи оценки многомерной плотности к получению оценок одномерных плотностей и обращению преобразования Радона. Нетрудно убедиться, что справедлива следующая теорема [2].

**Теорема** Пусть  $X$  –  $d$ -мерный случайный вектор с плотностью  $f(x)$ . Тогда случайная величина  $\langle X, \theta \rangle$  имеет плотность  $R_{\theta}f[s]$ .

Эта теорема дает следующий алгоритм решения рассматриваемой задачи. Спроецируем наблюдения  $X_1, X_2, \dots, X_N$  на некоторые направления  $\theta_1, \theta_2, \dots, \theta_M$ , вычислим оценки для  $R_{\theta_i}f[s]$  согласно формуле (1). Оценку для исходной плотности  $f(x)$  можно получить с помощью методов обращения преобразования Радона [3].

## 4. Применение метода

В рамках данного исследования была произведена реализация вышеописанного метода. Для проверки работоспособности метода использовалась генерация случайных векторов, распределенных с заведомо известными плотностями. Оценки строились на основе 500 наблюдений. Таким образом, можно сделать вывод о пригодности метода для решения данной задачи путем сравнения полученных по методу оценок с модельными плотностями.

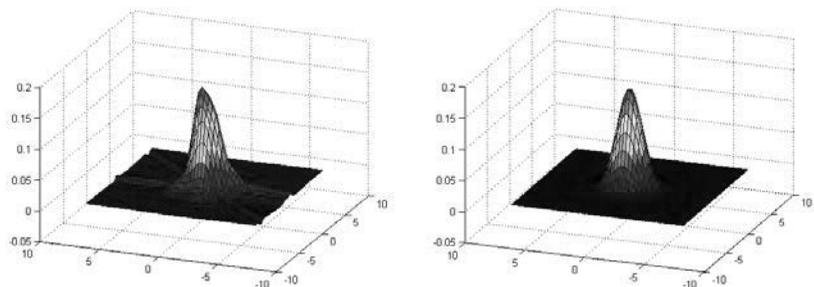
При реализации метода использовались вейвлеты Добеши 5 порядка. Для построения графиков, представленных ниже, был использован нелинейный вариант формулы (1) с жесткой пороговой обработкой

$$f(x) = \sum_k c_{j_0,k} \varphi_{j_0,k}(x) + \sum_{j_0}^{j_1} \sum_k d_{j,k} I(|d_{j,k}| > \delta) \psi_{j,k}(x),$$

где  $\delta$  – некоторый порог. Подробнее методы выбора порога и другие виды пороговой обработки изложены, например, в [4].

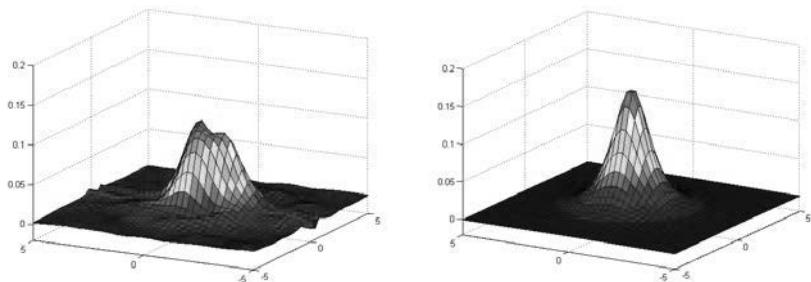
Метод показал неплохие результаты для таких распределений, как двумерное равномерное распределение в заданном квадрате, двумерное нормальное распределение и двумерное распределение Стьюдента. Результаты работы метода для нормально распределенных наблюдений представлены на Рис.1. Слева – график плотности, полученный по методу, справа – график плотности двумерного стандартного нормального распределения. Как видим, полученные графики имеют

визуальное сходство.



**Рис. 1.** Результаты для нормального распределения

Аналогичная процедура построения графиков была проделана и для стандартного двумерного распределения Стьюдента. Результаты представлены на Рис. 2.



**Рис. 2.** Результаты для распределения Стьюдента

Стоит отметить, что метод действует и для более высоких размерностей случайных векторов, чем 2, однако для размерности 3 и более визуализировать полученные результаты не представляется возможным.

## **Выводы**

Вышеописанный метод оценки плотностей дает возмож-

ность оценивать плотность распределения, не имея практически никаких представлений о ее свойствах и структуре. Такой способ представляется перспективным для дальнейшего изучения, поскольку локальная природа вейвлетов, скорее всего, позволит получать сравнительно хорошие оценки плотностей, обладающих, к примеру, резкими локальными осцилляциями и не обладающих свойством гладкости.

### **Благодарности**

*Выражаю благодарность своему научному руководителю Шестакову Олегу Владимировичу за предложение исследовать вышеописанный метод и помощь в поиске необходимой литературы.*

### **Список литературы**

1. *Vidakovic B.* Statistical Modeling by Wavelets – NY: John Wiley & Sons, 1999. – 408 p.
2. *O'Sullivan F., Pawitan Y.* Multidimensional Density Estimation by Tomography // Journal of the Royal Statistical Society. Series B (Methodological), 1993. Vol. 55. No. 2. P. 509–521.
3. *Шестаков О. В.* Непараметрическое оценивание многомерной плотности с помощью вейвлет-оценок одномерных проекций // Информатика и ее применения, 2015. Т.9. No.2. С. 88–92.
4. *Захарова Т. В., Шестаков О. В.* Вейвлет-анализ и его приложения. – М.: МАКС Пресс Москва, 2009. – 152 с.

# Имитационная модель буферизации видеопотока в одноранговой сети с учетом геолокации и активности пользователей\*

И.Ю. Васильев<sup>11</sup>, Ю.В. Гайдамака<sup>12</sup>

**Аннотация:** Построена имитационная модель буферизации видеопотока в пиринговых сетях с учетом геолокации пользователей, задержки данных между ними, стратегии загрузки и других особенностей процесса обмена видеоданными в P2P-сетях.

**Ключевые слова:** P2P-технологии; имитационная модель; одноранговые сети

## Введение

Технология P2P используется крупнейшими поставщиками услуг на рынке онлайн-телевидения, такими как BBC iPlayer, Zattoo, PeerCast, Pulse, QQLive и многими другими. Реализация концепции P2P-технологии в реальных сетях позволила свести к минимуму затраты коммерческих компаний на техническое оборудование. Эффективно используя ресурсы пользователей сети, P2P-технология минимизирует нагрузку на сервер-первоисточник видеоконтента. Таким образом, отпадает необходимость в установке дополнительных серверов и увеличении пропускной способности сети, чтобы справиться с большим количеством клиентов. Другим плюсом внедрения P2P-технологии является высокая отказоустойчивость сети: достаточно обеспечить бесперебойную работу хотя бы одного сервера-первоисточника видеоданных, чтобы гарантировать предоставление услуг всем пользователям на допустимом

---

\*Работа выполнена при поддержке РФФИ в рамках научных проектов № 14-07-00090, 15-07-03608, 15-07-03051.

<sup>11</sup> РУДН; 1992 г.р.; iuvasiliev@gmail.com

<sup>12</sup> РУДН; к.ф.-м.н.; ygaidamaka@mail.ru

уровне QoE.

Однако у P2P-сетей есть и недостатки. Среди основных – недостаточная информационная безопасность пользователей, сравнительно большая задержка данных, а также необходимость пользователей быть в постоянном поиске недостающих фрагментов видео. Если недостаток информационной безопасности – это вынужденная жертва в погоне за прибылью, то большую задержку видеопотока и коррелирующий с ней недостаток данных в сети можно и нужно минимизировать. Это цели служат разрабатываемые модели P2P-сети, как математические, так и имитационные.

Известны модели P2P-сети с учетом так называемых лагов – задержек передачи данных от сервера и между пользователями. В наших предыдущих работах [1-4] эти модели позволили сформулировать и получить решения задачи оптимизации выбора наборов действующих в сети стратегий. На основе этих работ разработана новая модель функционирования потоковой P2P-сети, в которой учитывается географическое местоположение каждого пользователя, удаленность пользователей от сервера и друг от друга, сложность маршрутовк пользователям и их суточная активность.

## **1. Имитационная модель**

В статье построена имитационная модель потоковой P2P-сети с одним сервером-первоисточником видеоданных и N разнородными пользователями, просматривающими видео в режиме реального времени. Видео делится на равные порции, небольшие фрагменты, передача которых и образует поток видеоданных. Время моделирования разбито на такты. Длина такта постоянна и равна времени воспроизведения одной порции данных. Для простоты будем считать, что один такт равен одной секунде. Таким образом, мы представляем одни астрономические сутки в виде  $24 \times 60 \text{ мин} \times 60 \text{ с} = 86\,400$  тактов.

Количество пользователей в сети не постоянно. Каждый пользователь за сутки пребывает в сети случайное время, рас-

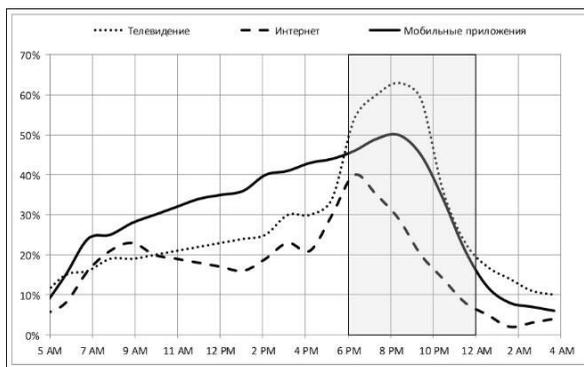
пределенное по экспоненциальному закону со средним значением, равным 21 600 тактам (6 часов):

$$HoursOnline(n) \sim P(21600), n = \overline{1, N}.$$

Введем параметр активности пользователя, который отражает случайный характер поведения пользователей:

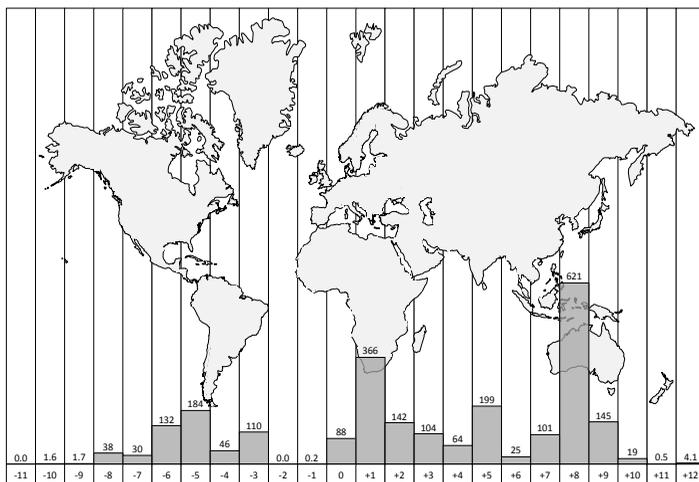
$$UserActivity(n) \sim RAND(1..100), n = \overline{1, N}$$

Эта величина показывает, как часто пользователь приходит в сеть, уходит из нее и как часто переключается с канала на канал. Здесь  $UserActivity(n) = 1$  означает, что в течение суток пользователь один раз пришел в сеть и находился в ней случайное время, имеющее пуассоновское распределение с параметром  $HoursOnline(n)$ , не переключаясь на другие каналы.  $UserActivity(n) = i$  означает, что  $n$ -пользователь подключался к сети  $i$  раз за сутки, причем, каждый из сеансов длился ровно  $HoursOnline(n)/i$  тактов. Таким образом, если  $UserActivity(n) = 100$  и  $HoursOnline(n) = 15000$  тактов, пользователь проведет за сутки 100 сеансов по 150 тактов каждый. На Рис. 1 представлено распределение количества пользователей в сети в зависимости от времени суток. [7]



**Рис. 1.** Распределение количества пользователей в сети

В соответствии с этим распределением и значениями  $UserActivity = (UserActivity(n))_{n=1, N}$  и  $HoursOnline = (HoursOnline(n))_{n=1, N}$ , случайным образом генерируются интервалы пребывания пользователя в сети  $TimeOnline = (TimeOnline(n))_{n=1, N}$  – бинарная матрица размера  $N \times T$ . Матрица показывает такты, на которых пользователи находятся онлайн:  $TimeOnline(n, l) = 1$ , если  $n$ -пользователь на  $l$ -такте находится в сети, и  $TimeOnline(n, l) = 0$  в противном случае. Таким образом, если  $TimeOnline(n, l) = 1$  и  $TimeOnline(n, l + 1) = 0$ , то  $n$ -пользователь покинул сеть на  $(l + 1)$ -ом такте, а значения  $TimeOnline(n, l) = 0$  и  $TimeOnline(n, l + 1) = 1$  говорят о том, что на  $(l + 1)$ -м такте  $n$ -пользователь вернулся в сеть.



**Рис. 2.** Распределение пользователей по часовым поясам

Приход новых пользователей в сеть и их уход из сети существенно влияют на ключевые показатели эффективности. Так, когда новый пользователь подключается к сети, у него еще нет данных для обмена с другими пользователями, зато он использует ресурсы других пользователей, чтобы загрузить себе контент. В то же время, когда пользователь уходит из сети, он перестает участвовать в раздаче уже закачанных порций видео.

Все  $N$  пользователей сети разбиты случайным образом по часовым поясам в соответствии с распределением, показанным на Рис. 2 [8].

В модели разбиение пользователей по часовым поясам позволяет уменьшить вероятность глобальных всплесков, соответствующих массовым отключениям пользователей от сети и массовым подключениям.

Чтобы сгладить паузы в воспроизведении, а также синхронизировать видео и звук, на стороне пользователя имеется буфер данных для загрузки и хранения  $M+1$  порции данных, где 0-место предназначено для загрузки новых порций данных от сервера-первоисточника, а оставшиеся  $M$  мест — для хранения и обмена порциями с другими пользователями. Таким образом, состояние  $n$ -пользователя представимо в виде вектора  $x(n) = (x_0(n), \dots, x_M(n))$  где  $x_m(n) = 1$ , если  $n$ -пользователь имеет порцию данных на  $m$ -позиции буфера, и  $x_m(n) = 0$  в противном случае. [1-6]

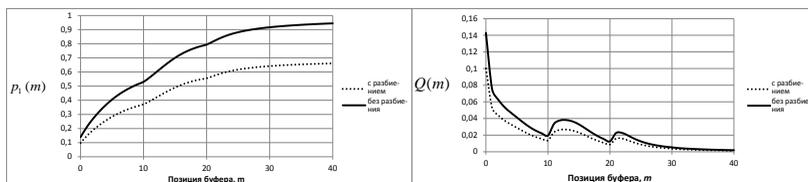
Лаг – это количество тактов между отправкой порции данных и ее получением, таким образом, лаги отражают количественную характеристику задержки порции данных. Так как P2P-сеть является наложенной, все характеристики соединения у пользователей сохраняются. Это означает, что каждый пользователь обладает уникальными скоростью загрузки, скоростью отдачи, величиной задержки данных и набором соседей (пользователей, с которыми происходит взаимодействие напрямую). Уникальные наборы этих характеристик делают пользователей P2P-сети разнородными.

В процессе потокового вещания в P2P-сети действуют различные заранее определенные стратегии. Всего их три: стратегия составления списка соседей, стратегия выбора целевого пользователя из списка соседей и стратегия выбора порции данных для загрузки. Выбор стратегий влияет на основные показатели функционирования P2P-сети: вероятность непрерывного воспроизведения видео, задержку начала воспроизведения, вероятность наличия и загрузки порции данных на позициях буфера, а также вероятность возникновения коллизии –

ситуации, когда запрашивающий пользователь не сможет за- качать порцию данных из-за того, что у отдающего пользова- теля не хватает пропускной способности.

## 2. Численный эксперимент

Задача численного эксперимента – сравнить результаты мо- дели, не учитывающей распределение пользователей по вре- менным зонам [1-6], с новой моделью. В качестве исходных данных были выбраны  $N = 300, M = 40$ . Время моделирования  $T = 1\,000\,000$  тактов (около 12 суток модельного времени). За- держки у пользователей равны  $lag = \{0,10,20\}$ . Стратегией за- грузки видеоданных была выбрана Rarest First. Соседи, как и целевой пользователь, выбирались равновероятно.



**Рис. 3.** Вероятность наличия  $p_1(m)$  и вероятность загрузки порции данных  $Q(m)$  на позиции буфера.

Как видно из Рис. 3, результаты моделирования потоко- вой P2P-сети с разбиением пользователей по временным зонам (пунктирная линия) качественно повторяют результаты моде- лирования с однородными по времени пользователями (сплош- ная линия). Характер поведения кривых, в том числе перегибы в точках  $m = 10$  и  $m = 20$ , подробно описан в [1-5]. Значительно более низкая вероятность непрерывного воспроизведения ви- део  $p_1(40)$  для модели с временными зонами объясняется тем, что доставка порций данных от удаленных соседей может зани- мать более одного такта. Результаты показывают, что сети с за- данным набором исходных данных требуются совершенство- вание механизма распространения видео.

## Выводы

Целью дальнейших исследований является постановка

и аналитическое решение задачи оптимизации обмена данными в одноранговой сети. В ближайших планах расширение имитационной модели за счет модификации стратегий загрузки порций данных.

## Список литературы

1. *Gaidamaka Yu., Vasiliev I., Samuylov A., Samuylov K., Shorgin S.* Simulation of Buffering Mechanism for Peer-to-Peer Live Streaming Network with Collisions and Playback Lags // In Proc. of the 13th ICN, 2014, pp. 86-91.
2. *Gaidamaka Yu., Samuylov A., Medvedeva E., Vasiliev I., Abaev P., Shorgin S.* Design and Software Architecture of Buffering Mechanism for Peer-To-Peer Streaming Network Simulation // Proc. 29th Europ. Conf. on Modelling and Simulation (ECMS), 2015, pp. 682-688.
3. *Гайдамака Ю.В., Самуйлов А.К., Васильев И.Ю.* Симулятор процесса буферизации при обмене данными в потоковой одноранговой сети // Тр. IX Межд. отраслевой научно-техн. конф. «Технологии информационного общества», Москва, МТУСИ, 2015, с. 13-20.
4. *Гайдамака Ю.В., Васильев И.Ю., Магамадов Р.С.* Имитационная модель оптимизации стратегии выбора соседей в потоковой одноранговой сети // Тр. Всеросс. конф. «Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем». Москва, РУДН, 2015, с. 85-87.
5. *Adamu A., Gaidamaka Yu., Samuylov A.* Discrete Markov Chain Model for Analyzing Probability Measures of P2P Streaming Network // LNCS, 2011, №.6869. pp. 428-439.
6. *Zhao, Y., Shen, H.* A simple analysis on P2P streaming with peer playback lags // Proc. of the 3rd Int. Conf. on Communication Software and Networks (IEEE ICCSN), 2011, pp. 396-400.
7. [http://www.phonearena.com/news/Study-shows-that-in-app-advertising-is-more-effective-than-television-or-internet-ads\\_id22556](http://www.phonearena.com/news/Study-shows-that-in-app-advertising-is-more-effective-than-television-or-internet-ads_id22556) Дата обращения: 26.09.2015г.
8. <http://royal.pingdom.com/2013/02/12/internet-users-time-zone-2/> Дата обращения: 26.09.2015г.

# Анализ естественных и искусственных изображений в рекламе посредством непараметрической оценки функции распределения плотности пикселей

Э.М. Гиоргис<sup>13</sup>

**Аннотация:** Данная работа является исследованием в области анализа изображений. С помощью методов сегментации анализируются изображения, типичные для рекламы, с объектом и фоном. Целью исследования являлось нахождение способа отличия естественного изображения от искусственного и применение его к рекламным изображениям.

**Ключевые слова:** анализ изображений, реклама, синтетические изображения

## Введение

Изображения являются одним из распространенных способов представления информации. С развитием цифровых технологий изображения получили широкое применение. Графический тип рекламы является основным объектом, исследуемым в данной работе. Первым этапом взаимодействия рекламы с потребителем является привлечение внимания [1,2,3]. При этом реклама становится агрессивней, что приводит к негативному отношению к ней потребителей и, как следствие, уменьшению внимания. Поэтому в последнее время появляется так называемая нативная реклама, призванная быть ненавязчивой. Примером можно назвать статью в печатном издании, спонсируемую брендом, поисковую рекламу. Кроме того, появ-

---

<sup>13</sup> МФТИ; 1994 г.р.; giorgis@phystech.edu

ляются технологичные методы создания рекламы. Так, например, в 2010 году Мерседес-Бенц успешно применил гибридные изображения в своей наружной рекламе [4].

Разработка методов создания нативной рекламы является актуальной, и один из таких методов предлагается в данной работе.

## 1. Используемые методы анализа изображения

### 1.1. Процедура сдвига среднего

Для выбора правильного метода сегментации для обработки рекламных изображений нужно учесть низкий приоритет точности и важность скорости и минимизации количества параметров сегментации. Поэтому в данной работе был выбран метод Mean-shift [5], основанный на методе окна Парзена [6]. Оценка плотности распределения с помощью ядра  $K$ , его профиля  $k$ , константой нормирования  $c_{k,d}$  (или метод окна Парзена с шириной  $h$ ) определяется формулой (1).

$$\hat{f}_{h,K}(\mathbf{x}) = \frac{c_{k,d}}{nh^d} \sum_{i=1}^n k\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right) \quad (1)$$

Логика данного подхода состоит в рассмотрении пространства характеристик как эмпирической функции распределения вероятности представленного параметра. Более плотные области в пространстве состояний соответствуют локальным максимумам функции распределения вероятности, т.е. модам этого неизвестного распределения. Если определено местоположение моды, устанавливаются очертания связанного с ней кластера [7].

Первый шаг анализа пространства характеристик с распределением  $f(\mathbf{x})$  состоит в нахождении мод плотности. Эти моды локализуются среди нулей градиента  $\nabla f(\mathbf{x})=0$  и процедура

сдвига среднего – это способ локализовать эти нули, не оценивая распределение вероятности.

Определим функцию  $g(x)$  как (2) и, используя ее для обозначения профиля, определим ядро  $G(x)$  как (3). Тогда, используя оценку градиента распределения, получим выражения для вектора сдвига среднего (4).

$$g(x) = -k'(x) \quad (2)$$

$$G(x) = c_{g,d} g(\|x\|^2), \text{ где } c_{g,d} - \text{константа нормирова} \quad (3)$$

$$\mathbf{m}_{h,G}(\mathbf{x}) = \frac{\sum_{i=1}^n x_i g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right)} - \mathbf{x} = \frac{1}{2} h^2 c \frac{\hat{\nabla} f_{h,K}(\mathbf{x})}{\hat{f}_{h,G}(\mathbf{x})} \quad (4)$$

Таким образом, вектор сдвига среднего всегда направлен в одну сторону с градиентом, то есть к моде распределения. Алгоритм состоит в циклическом повторении (пока вектор сдвига отличен от нуля) построения вектора сдвига и переноса центра окна Парзена. Нулевой вектор сдвига означает, что текущая точка – локальный максимум распределения. Для сегментации посредством метода Mean-shift был использован программный продукт EDISON.

## 1.2. Слияние сегментов

Рекламные изображения содержат в себе изображение объекта и фон, поэтому представляется полезным использование метода сегментации для выделения главного объекта на изображении. В качестве этого метода был выбран метод MSRM (Maximal-similarity based region merging) [8].

Ключевой вопрос при соединении сегментов состоит в том, как определить сходство непомеченных сегментов с уже помеченными сегментами, чтобы похожие сегменты могли соединяться автоматически. В работе [8] для определения сходства сегментов  $R$  и  $Q$  используется мера косинуса (коэффициент Бхаттачарая) (5).

$$\rho(R, Q) = \sum_{u=1}^{4096} \sqrt{Hist_R^u \cdot Hist_Q^u} \quad (5)$$

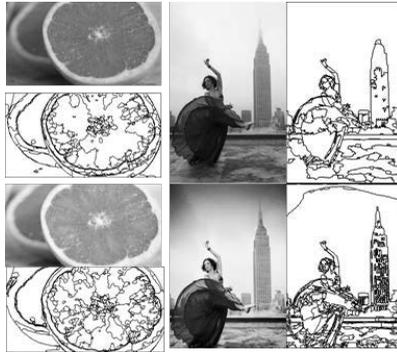
Здесь Hist – нормированная гистограмма сегмента, нижний индекс – сегмент, а верхний индекс – элемент. Чем больше мера косинуса между сегментами, тем выше их сходство.

## **2. Описание решения задачи нахождения правила отличия естественного изображения от искусственного**

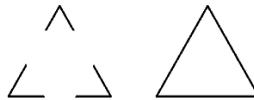
Для нахождения свойств естественных и искусственных изображений проведено 17 экспериментов (каждый эксперимент – анализ естественного изображения, его обработанной искусственной копии и их n-кратно увеличенных дубликатов, n=3, 6).

Было обнаружено, что на сегментированных естественных изображениях не прослеживаются многие четкие границы, соответствующие восприятию границ человеческим глазом, в то время как на искусственных изображениях эти границы имеют ярко выраженный характер (см. Рис. 1).

Границы на искусственном изображении согласуются с человеческим представлением об объекте. Простым примером, объясняющим факт такой схожести, является изображение трех уголков на вершинах невидимого треугольника. Человеческий мозг самостоятельно «дорисовывает» недостающие стороны (границы), так что три разрозненных линии воспринимаются как треугольник (см. Рис. 2).



**Рис. 1.** Эксперимент (границы на естественных и искусственных изображениях)



**Рис. 2.** Пример-треугольник

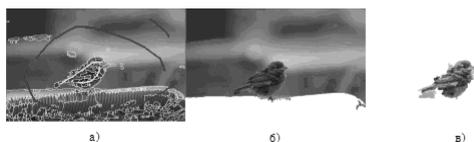
В большинстве случаев (в 14 из 17) наблюдалось увеличение количества сегментов после обработки изображения. Это можно объяснить тем, что в тех местах, где на естественных изображениях воспринималась граница, но отсутствовала в виде однородной линейки пикселей, на искусственных эта граница проявлялась после применения сегментации, увеличивая количество сегментов (см. Рис. 3).



**Рис. 3.** Схематическое изображение границ естественных и искусственных изображений

Это же наглядно подтверждается результатами исполь-

зования алгоритма MSRM. При обработке естественных изображений границы выделяемых объектов получались нечеткими, вследствие чего результат применения алгоритма MSRM к естественным изображениям не давал желаемых результатов (контур объекта не соответствовал контуру, представляемому человеком). Применение алгоритма MSRM к искусственным изображениям показало, что контур объекта оказывался ярко выраженным на любом уровне приближения (см. Рис. 4).



**Рис. 4.** Применение MSRM к естественному (б) и искусственному (в) изображениям. На оригинале (а) нанесены маркеры объекта и фона.

### 3. Описание результатов

В данной работе было исследовано 23 рекламных изображения из Интернета и 18, полученных самостоятельно, всего 41. Они также были разбиты на сегменты, на них были выделены границы. В результате этого у объектов на ряде изображений были обнаружены четкие границы, что соответствует тому, что собой представляет искусственное изображение. Кроме того, после применения алгоритма MSRM к рекламным изображениям, были получены четкие контуры объектов, представленных на рекламном изображении.

### Выводы

Обнаружено различие между естественными и искусственными изображениями в виде четких границ у объектов на искусственных изображениях. Показано увеличение количества сегментов при переходе от естественного изображения к

искусственному. Проведенный анализ результатов показал искусственный характер ряда рекламных изображений. Построенные инструменты позволяют создателям рекламы выполнять анализ изображений на естественность.

## Список литературы

1. *Wijaya, Bambang Sukma*. The Development of hierarchy of effects model in advertising / Wijaya, Bambang Sukma // International Research Journal of Business Studies, 2012. Vol.1. No.5. P. 73–85.
2. *Пронин С.* Рекламная иллюстрация. Креативное восприятие. – М.:Бератор-Пресс, 2003 – 144 с.
3. *Шугерман Д.* Искусство создания рекламных посланий. – М.: Манн, Иванов и Фербер, 2012. – 400 с.
4. *Aude Oliva, Antonio Torralba and Philippe G. Schyns*. Hybrid images // ACM Transactions on Graphics, 2006. Vol.3. No.25. P. 527–532.
5. *Y. Cheng*. Mean shift, mode seeking, and clustering // IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995. Vol.8. No.17. P. 790–799.
6. *K. Fukunaga, L. Hostetler*. The estimation of the gradient of a density function, with applications in pattern recognition // IEEE Transactions on Information Theory, 1975. V.1.No.21. P. 32–40.
7. *J.G. Touzani, J.G. Postaire*. Clustering by mode boundary detection // Pattern Recognition Letter, 1989. No.9. P. 1–12.
8. *Jifeng Ning et al*. Interactive Image Segmentation by Maximal Similarity Based Region Merging // Pattern Recognition, 2010. Vol.2. No.43. P. 445–456

# Опыт разработки и испытаний флуориметра для экспресс-оценки жизнеспособности бактериальных культур\*

П.В. Горбунов<sup>14</sup>

**Аннотация:** Рассматриваются вопросы разработки, развития и применения флуориметра в медико-экологических информационных технологиях, и результаты его испытаний.

**Ключевые слова:** флуориметр, флуоресценция, микробиология, оптико-электронный прибор

## Введение

Прибор предназначен для счета микроорганизмов (в образце жидкости), что позволяет использовать его как в науке, так и в задачах клинической и гигиенической медицины, для определения концентрации микроорганизмов в питьевой воде, сточных водах, биологических жидкостях (моча, кровь) и т.д., а также, при дополнительных действиях по подготовке пробы, в воздухе помещений (больниц, стерильных боксов), или на твердых поверхностях. Разработанный прибор можно условно разделить на два блока: оптический и электронный.

## 1. Описание прибора

### 1.1. Оптический блок флуориметра

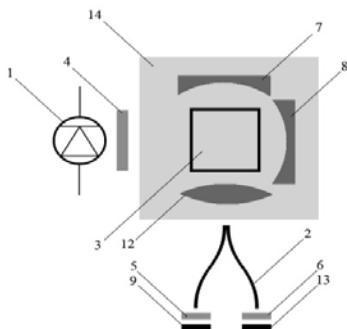
Оптический блок прибора (см. Рис. 1) собран по двухлучевой схеме и состоит из опорного и сигнального каналов. К оп-

---

\*Работа выполнена при поддержке коммерческой организации ООО «Биотермаль»

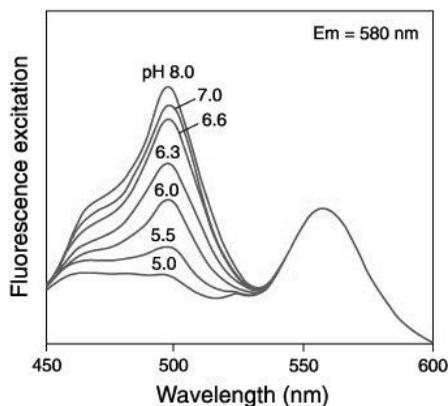
<sup>14</sup> МАИ НИУ; 1990; grbunv-pavel@rambler.ru

тическому блоку прибора относятся следующие элементы: синий светодиод (1), оптоволоконные линии (2), прямоугольная кювета с жидкой питательной средой и тестируемой культурой (3), заключенная в термостатируемую камеру (14), оптические фильтры (4,5,6), вогнутые сферические зеркала (7,8), фотоумножитель опорного канала (9), и фотоумножитель сигнального канала (13). Оптические фильтры (4,5), закрывающие светодиод и фотоумножитель опорного канала, имеют полосу пропускания 520-550 нм. Фильтр сигнального канала (6), пропускающий только флуоресцентное свечение краски, имеет полосу пропускания 480-500 нм. Прибор имеет две одинаковые измерительные ячейки.



**Рис. 1.** Схема оптического блока флуориметра

Принцип действия заключается в том, что в процессе деления микроорганизмов изменяется рН среды, а интенсивность свечения флуоресцентного красителя зависит от рН (см. Рис. 2), что и регистрирует прибор.



**Рис. 2.** Зависимость флуоресцентного выхода от длины волны при различных pH среды.

## 1.2. Электронный блок флуориметра

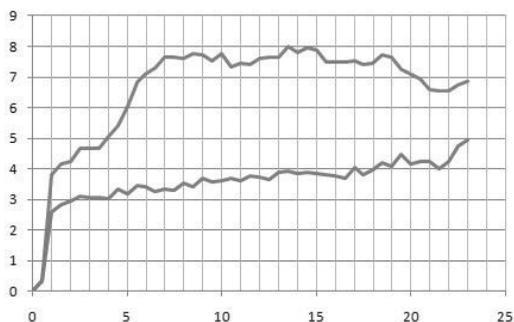
Электронный блок имеет два канала: амплитудный и фазовый. Для повышения чувствительности в блоке используются твердотельные фотоумножители. Свет от источника попадает на фотоумножители, а с выходов фотоумножителей попадает в амплитудный канал и фазовый канал. На выходе амплитудного канала формируется сигнал, пропорциональный логарифму отношения входных сигналов. На выходе фазового канала формируется сигнал, пропорциональный разности фаз входных сигналов. Затем полученные сигналы цифруются при помощи 10-битных АЦП, и запоминаются микропроцессором.

## 2. Испытания

Испытания проходили в два этапа. На первом этапе испытания были проведены на 20-ти образцах с различным количеством микроорганизмов (от  $10^3$  до  $10^6$  клеток). Использовалась стандартная стеклянная кювета размером 10x10x45 мм, кол-во питательной среды составляло 1мл. В каждый образец добавлялся флуоресцентный краситель флуоресцеин в количе-

стве 0,05%. Во время эксперимента с помощью термостата поддерживалась оптимальная для культивирования микроорганизмов температура в  $55 \pm 1^\circ\text{C}$ . Измерение флуоресценции проводилось с интервалом 30 мин в течение 10-24 часов. Для увеличения точности в каждой точке проводилось 1000 измерений. Так как прибор имеет две измерительные ячейки, за один эксперимент можно исследовать сразу два образца. Верхняя кривая показывает интенсивность свечения образца с клетками, а нижняя – стерильного образца.

На Рис. 3 изображен результат эксперимента с образцом, содержащим  $10^3$  микробных клеток, а на Рис. 4 –  $10^6$ . По горизонтальной оси отложено время в часах, а по вертикальной – светимость образца в условных единицах АЦП.

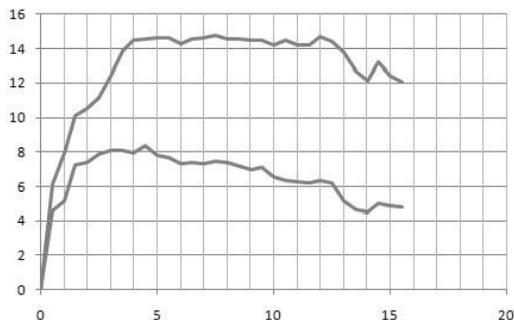


**Рис. 3.** Эксперимент с образцом, содержащим  $10^3$  клеток

Испытания первого этапа показывают работоспособность флуориметра и его пригодность для использования в микробиологических исследованиях. Однако, было выявлено несколько недостатков.

На Рис 3 и 4 видно, что в течение 30-50 минут после помещения образцов в прибор, когда происходит увеличение температуры образца от комнатной до  $55^\circ\text{C}$ , происходит увеличение интенсивности флуоресценции. Ионы флуоресцеина в растворе образуют ассоциации, либо вступая во взаимодействие с

компонентами питательной среды, либо образуя димеры. В таких ассоциациях происходит тушение флуоресценции. При повышении температуры ассоциации флуоресцеина разрушаются, тушение уменьшается, а интенсивность флуоресценции растет.



**Рис. 4.** Эксперимент с образцом, содержащим  $10^6$  клеток

Нижняя кривая (стерильный образец) представляет собой контроль, на фоне которого проявляются изменения в образце, где происходит рост культуры. Однако на рисунках видно, что в течение всего периода наблюдения происходит изменение сигнала от стерильного образца. Наиболее вероятной причиной является изменения напряжения питания и/или температуры твердотельных фотоумножителей.

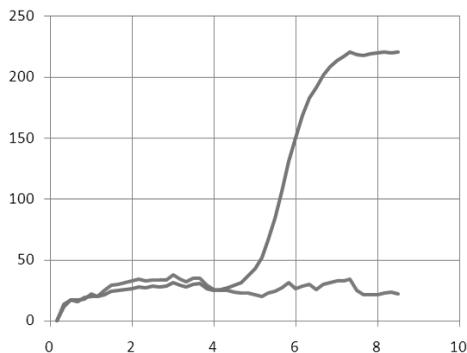
Выбранная разрядность АЦП недостаточна, так как полезный сигнал (разность между образцами) составляет 3 до 7 единиц.

Для проведения второго этапа испытаний были сделаны следующие изменения:

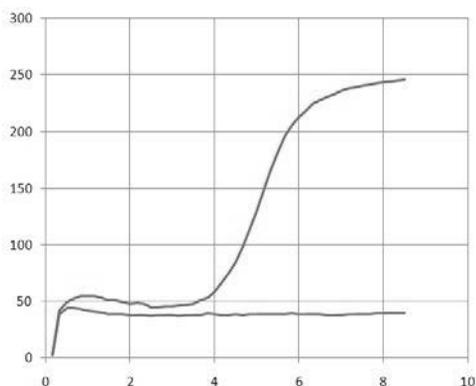
- изменен флуоресцентный краситель (пиранин);
- в электронный блок добавлены схемы стабилизации питания фотоумножителей и термостабилизации;
- увеличена разрядность АЦП с 10 до 12.

На Рис. 5 изображен результат эксперимента с образцом, содержащим  $10^3$  микробных клеток, а на Рис. 6 –  $10^5$ . По горизонтальной оси отложено время в часах, а по вертикальной

– светимость образца в условных единицах АЦП. Краситель – пиранин. Верхняя кривая – образец с клетками, нижняя – стерильный образец.



**Рис. 5.** Эксперимент с образцом, содержащим  $10^3$  клеток



**Рис. 6.** Эксперимент с образцом, содержащим  $10^5$  клеток

В экспериментах второго этапа удалось устранить выявленные ранее недостатки. Увеличение сигнала и уменьшение дрейфа сигнала от контрольного образца позволили существенно увеличить соотношение сигнал/шум.

## Выводы

Испытания показывают работоспособность флуориметра и хорошие перспективы его использования в микробиологических исследованиях, в пищевой промышленности и в санитарной медицине. Также, несомненно, к преимуществам прибора можно отнести его дешевизну по сравнению с импортными аналогами.

## Список литературы

1. Пугачев В. С., Синицын И. Н. Теория стохастических систем: Учеб. пособие. – М.: Логос, 2004. – 1000 с.
2. Синицын И. Н., Горбунов П. В. Оптико-электронный флуориметр для экспресс-оценки жизнеспособности бактериальных культур // Оптико-электронные приборы в системах распознавания образов, обработки изображений и символьной информации "Распознавание - 2015": сб. матер. XII междунар. науч.-техн. конференции; Юго-Зап. гос. ун-т. Курск, 2015. С. 333-336.
3. Синицын И. Н., Сергеев В. И., Синицын В. И., Корепанов Э. Р., Белоусов В. В., Горбунов П. В. Опыт синтеза дискретных фильтров Пугачева для обработки нормальных процессов в эрдитарных стохастических системах. // XVI международная научно-техническая конференция «Кибернетика и высокие технологии XXI века», Воронеж 2015. НПФ «Саквоее» ООО. С. 76–88. ISBN 978-5-904259-32-7.
4. Sinitsyn I. N., Gorbunov P. V. The experience in the developing suboptimal algorithm of analysis, modeling and interpreting signals in outer scanning analyzer of microorganisms. // 11th International Conference on Pattern Recognition and Image Analysis: New Information Technologies (PRIA-11-2013). Samara, September 23–28, 2013. Conference Proceedings (Vol. I-II), Volume II, Samara: IPSI RAS, 2013. З. 471–472
5. Синицын И. Н., Горбунов П. В. Условно оптимальные методы и средства совместного аналитического и статистического моделирования процессов и их применение в космическом сканирующем анализаторе микроорганизмов. // Оптико-электронные приборы в системах распознавания образов, обработки изображений и символьной информации «Распознавание – 2013»: сб. матер. XII междунар. науч.-техн. конференции; Юго-Зап. гос. ун-т. Курск, 2013. С. 426–428.

# Оценка времени отклика обработки запросов в системе облачных вычислений

А.В. Горбунова<sup>15</sup>

**Аннотация:** Облачные вычисления представляют собой технологию, позволяющую удаленному пользователю получать доступ к различным ресурсам через интернет. В статье рассматривается система облачных вычислений, в которую поступают сложные пользовательские запросы, состоящие из нескольких подзапросов. Для анализа одного из важнейших показателей качества функционирования такой системы – времени отклика – построена математическая модель, позволяющая получить аналитическую формулу для нахождения оценки среднего значения и дисперсии времени отклика системы, т. е. максимального времени обработки составляющих запрос подзапросов. Для проверки точности приближения проведено имитационное моделирование средствами GPSS (General Purpose Simulation System).

**Ключевые слова:** система облачных вычислений; время отклика; система массового обслуживания

## Введение

Облачные вычисления или cloud computing представляют собой технологию, позволяющую удаленному пользователю в режиме реального времени по требованию получать доступ к вычислительным ресурсам через Интернет. В задачи поставщика таких услуг входит обеспечение требуемого уровня оказываемых услуг, а также избежание перегрузки ресурсов при обработке пользовательских требований [1–3].

Рассмотрим облачный центр, состоящий из нескольких физических машин, которые резервируются пользователями в

---

<sup>15</sup> РУДН; 1986 г.р.; avgorbunova@rambler.ru

порядке поступления запросов и могут использоваться совместно для их обработки посредством метода виртуализации. В систему поступают сложные запросы пользователей, содержащие несколько задач, для обработки каждой из которых требуется одна виртуальная машина. Запрос считается выполненным, т. е. пользователь получает отклик системы, после обработки всех его подзапросов. Таким образом, время отклика системы облачных вычислений определяется как максимальное время обработки составляющих его заданий или подзапросов [4, 5].

Основой для построения математической модели оценки времени отклика послужила так называемая fork-join система (система с расщеплением). Из-за существующей зависимости между очередями подзапросов вследствие общих моментов поступления, на сегодняшний день точные аналитические результаты для среднего времени отклика существуют только для случая расщепления на два подзапроса при ограничении равенства интенсивностей обслуживания [6]. Подробный обзор существующих публикаций по данной тематике представлен в [7].

## 1. Математическая модель

Рассмотрим систему массового обслуживания с  $K$  независимыми параллельно работающими приборами, каждый из которых имеет накопитель неограниченной емкости. В эту систему поступает пуассоновский поток заявок или запросов с интенсивностью  $\lambda$ . При этом в момент поступления заявка делится на  $K$  подзапросов (от 1-го до  $k$ -го типа), которые поступают в каждую из  $K$  очередей соответственно. Заявка считается обслуженной в момент окончания обслуживания последнего из составляющих ее подзапросов. Времена обслуживания заявок подзапросов являются независимыми случайными величинами, распределенными по экспоненциальному закону с параметром  $\mu_k$  для подзапроса  $k$ -го типа,  $k = \overline{1, K}$ .

Пусть  $x_k(t)$  – число подзапросов  $k$ -го типа, находящихся

в системе в момент времени  $t$ . Обозначим  $X(t) = (x_1(t), \dots, x_K(t))$ . Случайный процесс  $\{X(t), t \geq 0\}$ , описывающий поведение системы во времени, является марковским (МП) с множеством состояний  $X = \{\vec{n} := (n_1, \dots, n_k, \dots, n_K), n_1 \geq 0, \dots, n_k \geq 0, \dots, n_K \geq 0\}$ .

В предположении о его существовании стационарное распределение удовлетворяет следующей системе уравнений равновесия (СУР):

$$\left( \lambda + \sum_{k=1}^K \mu_k u(n_k) \right) p(\vec{n}) = \lambda p(\vec{n} - \vec{1}) \prod_{k=1}^K u(n_k) + \sum_{k=1}^K \mu_k p(\vec{n} - \vec{e}_k),$$

$$\vec{n} \in X,$$

где  $\vec{e}_k$  – вектор, все элементы которого равны нулю, кроме  $k$ -го, который равен 1,  $\vec{1}$  – единичный вектор, и

$$u(n_k) = \begin{cases} 0, & \text{если } n_k = 0, \\ 1, & \text{если } n_k > 0. \end{cases}$$

Здесь не исследуется очевидное условие эргодичности рассматриваемой системы и считается далее, что  $\rho = \max_{1 \leq k \leq K} \rho_k$ , где  $\rho_k = \lambda / \mu_k, k = \overline{1, K}$ .

## 2. Оценка времени отклика

Исследуя СУР, замечаем, что суммированием ее уравнений по всем индексам, кроме  $k$ , приходим к СУР, соответствующей системе массового обслуживания  $M/M/1$ . Следовательно, можем получить выражения для маргинальных вероятностей макросостояний МП  $X(t)$  в виде

$$p_{n_k} = (1 - \rho_k) \cdot \rho_k^{n_k}, n_k \geq 0.$$

При допущении о независимости времен пребывания подзапросов в подсистемах и с учетом того, что время отклика является максимумом времен пребывания составляющих заявку подзапросов, вычисляем функцию распределения времени отклика, как произведение соответствующих функций распределения для времен пребывания каждого из подзапросов в системе. Тогда математическое ожидание и дисперсия времени отклика  $\eta$  имеют вид

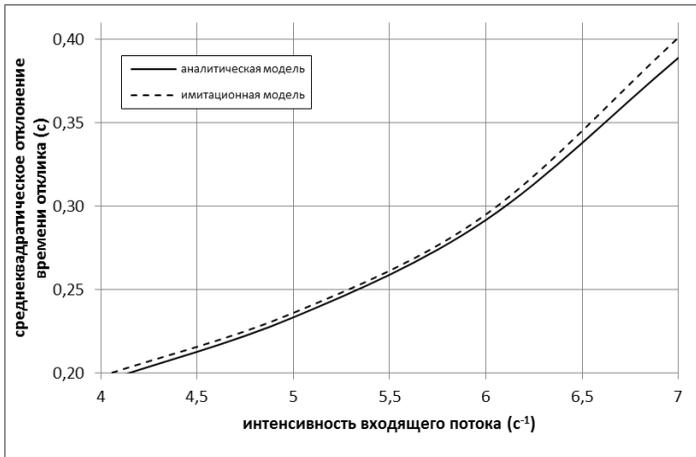
$$\begin{aligned}
E\eta &= \sum_{l=1}^K \frac{1}{\mu_l - \lambda} - \sum_{l \neq m} \frac{1}{\mu_l + \mu_m - 2\lambda} \\
&\quad + \sum_{l \neq m \neq n} \frac{1}{\mu_l + \mu_m + \mu_n - 3\lambda} + \dots + \\
&\quad + (-1)^{2K-1} \frac{1}{\mu_1 + \mu_2 + \dots + \mu_K - K\lambda}, \\
D\eta &= \sum_{l=1}^K \frac{2}{(\mu_l - \lambda)^2} - \sum_{l \neq m} \frac{2}{(\mu_l + \mu_m - 2\lambda)^2} + \\
&\quad + \sum_{l \neq m \neq n} \frac{2}{(\mu_l + \mu_m + \mu_n - 3\lambda)^2} + \dots + \\
&\quad + (-1)^{2K-1} \frac{2}{(\mu_1 + \mu_2 + \dots + \mu_K - K\lambda)^2} - \\
&\quad - \left( \sum_{l=1}^K \frac{1}{\mu_l - \lambda} - \sum_{l \neq m} \frac{1}{\mu_l + \mu_m - 2\lambda} + \right. \\
&\quad + \sum_{l \neq m \neq n} \frac{1}{\mu_l + \mu_m + \mu_n - 3\lambda} + \dots + \\
&\quad \left. + (-1)^{2K-1} \frac{1}{\mu_1 + \mu_2 + \dots + \mu_K - K\lambda} \right)^2.
\end{aligned}$$

Кроме того, можно записать выражения и для других характеристик системы массового обслуживания, используя аналогичные результаты для системы типа  $M/M/1$ .

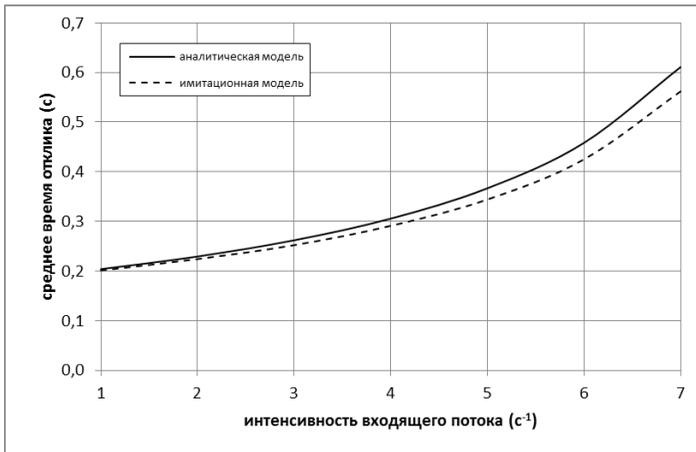
### 3. Численный пример

Для оценки полученной аппроксимации с помощью программных средств в системе имитационного моделирования GPSS была построена имитационная модель. Поскольку интерес представляют характеристики системы в стационарном режиме ее работы, имитационное моделирование проводилось до тех пор, пока не стабилизировались параметры модели. Стабилизация наступила после прохождения через систему около

500 000 запросов.



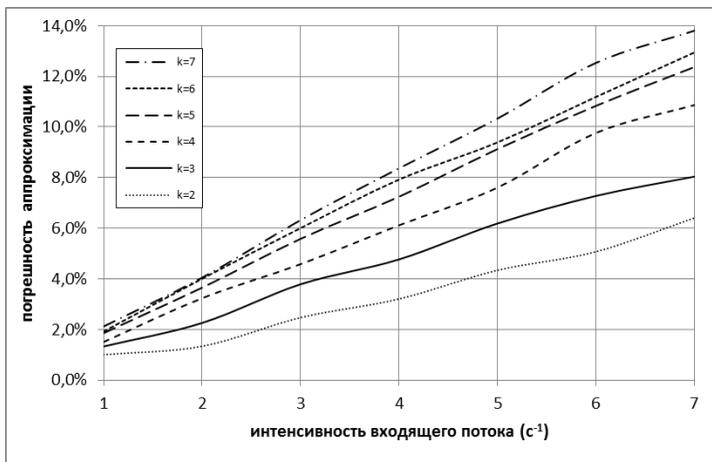
**Рис.1.** Среднее время отклика



**Рис.2.** Среднеквадратическое отклонение времени отклика

На Рис. 1 и 2 представлены графики зависимости среднего и среднеквадратического отклонения времени отклика системы облачных вычислений от интенсивности  $\lambda$  поступления запросов. Расчеты были выполнены для случая расщепления пользовательского запроса на 3 подзапроса ( $K = 3$ ) для симметричного случая  $\mu_1 = \mu_2 = \mu_3 = 10 \text{ с}^{-1}$  при меняющейся интен-

сивности входящего потока  $\lambda$ . Видно, что полученная с помощью аналитической модели оценка среднего времени отклика является оценкой «сверху», а среднеквадратического отклонения – оценкой «снизу». Заметим, что для несимметричного случая качественных изменений численных результатов не наблюдается.



**Рис.3.** Погрешность аппроксимации для среднего времени отклика

На Рис. 3 изображены графики для погрешности аппроксимации математического ожидания времени отклика в зависимости от интенсивности  $\lambda$  входящего потока для различных значений  $K = \overline{2,7}$ . Видно, что с увеличением числа подзапросов и интенсивности входящего потока увеличивается ошибка приближения.

## Выводы

В статье рассмотрена и проанализирована модель системы облачных вычислений, в которую поступают сложные запросы. Получена формула для расчета приближения времени отклика системы, реализована имитационная модель и прове-

ден сравнительный анализ. В дальнейшем планируется улучшить аналитические и имитационные результаты, а также исследовать модели с другим типом входящего потока и иными законами распределения времени обслуживания, провести расчет их основных вероятностно-временных характеристик и показателей качества обслуживания.

## Список литературы

1. *Xiong K., Perros H.* Service performance and analysis in cloud computing // Proceedings of the 5th IEEE World Congress on Services – Los Angeles: IEEE, 2009. P. 693–700.
2. *Satyanarayana A., Suresh Varma P., Rama Sundari M. V., Sarada Varma P.* Performance Analysis of Cloud Computing under Non Homogeneous Conditions // International Journal of Advanced Research in Computer Science and Software Engineering, 2013. Vol.3. P. 969–974.
3. *Buyya R., Broberg J., Goscinski A. M.* Introduction to cloud computing // Cloud Computing: Principles and Paradigms. – John Wiley & Sons, 2011. P. 3–42.
4. *Мокров Е. В., Самуйлов К. Е.* Модель системы облачных вычислений в виде системы массового обслуживания с несколькими очередями и с групповым поступлением заявок // Т-сomm – Телекоммуникации и транспорт, 2013. Т.7. No 11. С. 139–141.
5. *Basharin G. P., Gaidamaka Yu. V., Samouylov K. E.* Mathematical Theory of Teletraffic and Its Application to the Analysis of Multiservice Communication of Next Generation Networks // Automatic Control and Computer Sciences, 2013. Vol.47. No.2. P. 62–69.
6. *Nelson R., Tantawi A. N.* Approximate analysis of fork/join synchronization in parallel queues // IEEE Transactions on Computers, 1988. Vol. 37. P. 739–743.
7. *Thomasian A.* Analysis of fork-join and related queueing systems // ACM Comput. Surv. (CSUR), 2014. Vol.47. No.2. P. 1–71.

# Алгоритмы поиска арбитражных возможностей на валютном рынке

В. И. Гусынин<sup>16</sup>

**Аннотация:** Нахождение оптимального и эффективного алгоритма поиска арбитражных возможностей на валютном рынке представляет собой одну из основных задач алгоритмического трейдинга. В работе представлена модель поиска арбитражных возможностей с помощью алгоритмов обнаружения отрицательных циклов в орграфах.

**Ключевые слова:** поиск арбитражных возможностей; орграф; отрицательный цикл

## Введение

На современных российских валютных рынках ежедневно совершаются десятки, а иногда и сотни сделок, курсы валют меняются быстро и непредсказуемо. Под арбитражной возможностью принято понимать возможность извлечения прибыли посредством совершения нескольких логически связанных сделок с одинаковыми или связанными активами в один и тот же момент времени на разных рынках, либо на одном и том же рынке в разные моменты времени. Основная цель данной работы – найти оптимальный алгоритм поиска арбитражных возможностей на валютном рынке [1]. Стоимость валют в настоящий момент времени на рынке можно представить в виде полного орграфа с взвешенными ребрами, где наименования валют – это вершины, а обменный курс – это ребра. Таким образом, возникает задача нахождения цикла в орграфе после прохождения которого, единица исходной валюты преобразуется в более, чем одну единицу. Иначе говоря, необходимо найти такой цикл в орграфе, произведение весов ребер которого будут

---

<sup>16</sup> РУДН; 1994 г.р.; vldraven@gmail.com

больше единицы. Подробное описание данной модели приведено в первом разделе работы, во втором разделе с использованием данных с реального валютного рынка и с помощью алгоритма произведен поиск арбитражных возможностей.

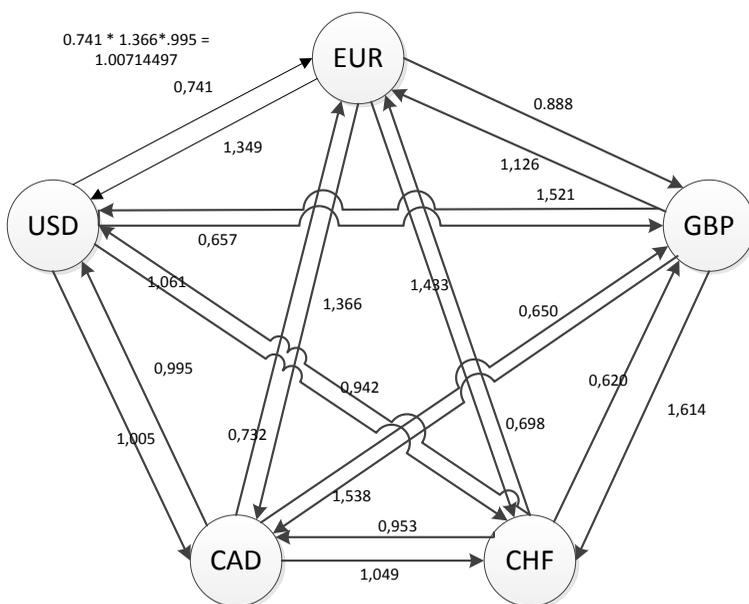
## 1. Описание модели

Если рассмотреть рынок финансовых транзакций, основанный на торговле валютой, то можно увидеть некоторые возможности применения задачи поиска отрицательных циклов. Если представить рынок финансовых транзакций, как в таблице 1, то по аналогии можно построить эквивалентный ей оргграф, показанный на Рис. 1.

Таблица 1. Обменный курс

	USD	EUR	GBP	CHF	CAD
USD	1	0,741	0,657	1,061	1,005
EUR	1,349	1	0,888	1,433	1,366
GBP	1,521	1,126	1	1,614	1,538
CHF	0,942	0,698	0,619	1	0,953
CAD	0,995	0,732	0,650	1,049	1

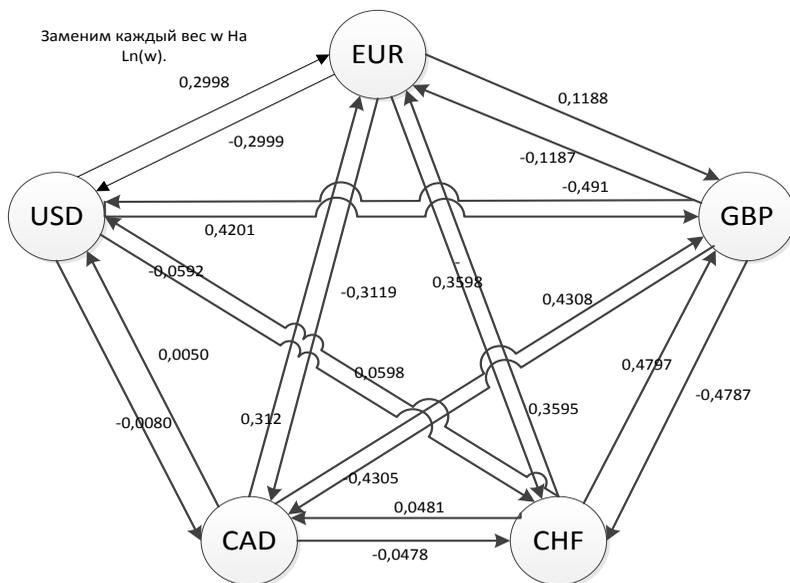
Для краткости здесь приведены только пять из сотен различных валют, торгуемых на современных валютных рынках: американские доллары (USD), евро (EUR), британские фунты (GBP), швейцарские франки (CHF) и канадские доллары (CAD),  $t$ -е число в  $s$ -й строке означает обменный курс - количество единиц валюты с именем в  $s$ -й строке, за которое можно купить 1 единицу валюты с именем в  $t$ -й строке. Например, из таблицы видно, что за 1000 американских долларов можно купить 741 евро.



**Рис. 1.** Орграф для таблицы 1

Эта таблица эквивалентна полному орграфу с взвешенными ребрами, где вершины соответствуют валютам, а ребра — обменным курсам. Ребро  $s \rightarrow t$  с весом  $x$  соответствует обмену  $s$  на  $t$  по цене  $x$ , а ребро  $t \rightarrow u$  с весом  $y$  соответствует обмену  $t$  на  $u$  по цене  $y$ . Пути на этом орграфе задают многошаговые обмены. Интересен случай, когда произведение весов ребер меньше веса ребра из последней вершины назад к первой. В нашем примере предположим, что вес ребра  $u \rightarrow s$  равен  $z$  и что  $xyz > 1$ . Тогда цикл  $s \rightarrow t \rightarrow u \rightarrow s$  позволяет обменять 1 единицу валюты  $s$  в более чем 1 единицу ( $xyz$ ) валюты  $s$ . То есть можно получить прибыль в  $100 \cdot (xyz - 1)$  процентов, поменяв  $s$  через  $t$  и и снова на  $s$ . Например, если обменять 1012,206 канадских долларов назад на американские доллары, мы получим  $1012,206 \cdot 0,995 = 1007,14497$  долларов — с прибылью в 7,14497 долларов. На первый взгляд, это немного, но валютный трейдер может оперировать миллионами долларов и может выполнять такие обмены каждую минуту — а это 7000 долларов в минуту или 420 000 долларов в час! Подобная ситуация представляет собой пример

возможности арбитража, который позволял бы трейдерам получать неограниченные прибыли, если бы не было ограничивающих факторов за пределами этой модели, таких как комиссия за транзакции или ограничения на размеры транзакций. Но даже и с учетом этих ограничений арбитраж может быть весьма прибылен в реальном мире.



**Рис. 2.** Орграф с логарифмированными весами ребер

Возникает вопрос: какое отношение имеет эта задача к кратчайшим путям? Ответ на этот вопрос довольно прост, задача арбитража представляет собой задачу обнаружения отрицательных циклов в орграфе с взвешенными ребрами. Данное утверждение доказывается следующим образом. Заменим каждый вес его логарифмом с обратным знаком. После этого вычисление весов путей с помощью перемножения весов ребер в исходной задаче соответствует их сложению в преобразованной задаче. А именно, любое произведение  $w_1 w_2 \dots w_k$  соответствует сумме  $-\ln(w_1) - \ln(w_2) - \dots - \ln(w_k)$ . Полученные веса ребер могут быть положительными или отрицательными, путь от  $v$  до  $w$

означает обмен валюты  $v$  на валюту  $w$ , а любой отрицательный цикл означает возможность арбитража.

## 2. Арбитраж на современных рынках

В этом разделе представлены результаты работы алгоритма на реальных данных с валютного рынка. Обменный курс занесен в Таблицу 2, аналогично Таблице 1. Данные получены с web-ресурса и представляют собой обменный курс в случайно выбранный момент времени. Разработанная на языке Java [2] программа Arbitrage.java использует алгоритм Беллмана Форда для обнаружения отрицательных циклов в заданном с помощью табличных значений орграфе, затем выводит на экран отрицательный цикл и возможную доходность.

**Таблица 2. Данные с реального рынка**

EUR	1,135	1	0,715	1,045	1,391
GBP	1,587	1,399	1	1,461	1,945
CHF	1,086	0,957	0,684	1	1,331
CAD	0,816	0,719	0,514	0,751	1

Результатом работы программы в приведенном примере являются следующие данные.

1000.00000 USD = 1226.00000 CAD

1226.00000 CAD = 1000.41600 USD

То, есть можно получить доход, совершая одновременно покупку и продажу CAD:

1000.00000 USD = 921.00000 CHF

921.00000 CHF = 1000.20600 USD

Также существует арбитражная возможность при покупке и продаже CHF:

1000.00000 EUR = 1391.00000 CAD

1391.00000 CAD = 714.97400 GBP

714.97400 GBP = 1000.24863 EUR

Также видно, что можно получить доход, проводя цикл

покупки CAD за EUR, затем покупки GBP за CAD, затем покупая EUR за GBP. Можно также привести пример более длинного цикла.

1000.00000 USD = 921.00000 CHF

921.00000 CHF = 881.39700 EUR

881.39700 EUR = 1226.02323 CAD

1226.02323 CAD = 630.17594 GBP

630.17594 GBP = 1000.08921 USD

## Выводы

На современных рынках роль информационных систем стала ключевой, это дало возможность многим разработчикам применять свои навыки в этой области. Основной задачей, которая остается так до конца и нерешенной, является нахождения оптимального алгоритма поиска арбитражных возможностей, т.е. алгоритма с минимальным временем выполнения. Также интересную и нерешенную задачу представляет собой алгоритм нахождения наиболее отрицательных циклов, алгоритм поиска наилучшей возможности арбитража. Сегодня многие крупные компании, которые занимаются сделкам и на валютной бирже использует торговых роботов, которые ежедневно и непрерывно проводят сделки покупки и продажи различных валют с целью извлечения прибыли.

Как и ожидалось, на современных рынках есть арбитражные возможности, в этом можно убедиться на примере полученных в ходе представленных исследований результатов.

## Список литературы

1. *Кац Д. О., Маккормик Д. Л.* Энциклопедия торговых стратегий. – М.: Альпина Паблишер, 2002. – 400 с.
2. *Седжвик Р., Уэйн К.* Алгоритмы на Java 4-е издание. – М.: ООО «ИД.Вильямс», 2013. – 848 с.

# Применение языка R для многомерной кластеризации временных рядов с целью анализа динамики научно-технического развития

Р.Д. Зайцев<sup>17</sup>, В.Б. Бритков<sup>18</sup>

**Аннотация:** В работе рассматривается задача многомерной кластеризации временных рядов на примере показателей динамики научно-технического развития стран мира. Предложен способ многомерного обобщения меры, основанной на коэффициенте временной корреляции CORT. Разработанный подход к решению описанной задачи реализован в среде статистического моделирования R.

**Ключевые слова:** научно-техническое развитие, многомерная кластеризация временных рядов, DTW, R.

## Введение

Анализ динамики и причинно-следственных связей научно-технического прогресса с темпами роста мировой экономики является примером такого типа исследований, интерес к которым за последние годы непрерывно возрастает. Однако, задача системной интеграции широкого набора методов, направленных на изучение данной проблемы, остаётся в достаточной степени нерешённой. В работе [1] отмечается, что общепринятые подходы содержат излишнее количество предварительных экстраполяционных гипотез, которые усложняют исследование глобальных процессов.

Для разработки непараметрических методов группи-

---

<sup>17</sup> МФТИ; 1992 г.р.; без учёной степени; romazaytc@mail.ru

<sup>18</sup> ФИЦ ИУ РАН (ИСА РАН); 1944 г.р.; к. ф.-м. н.; britkov@mail.ru

ровки динамик научно-технического развития стран мира в работах [2, 3] предложено применение инструментов кластеризации временных рядов с целью дальнейшего исследования полученных кластеров.. Наиболее удобными для решения подобного рода задач являются среды статистического моделирования, такие как Matlab, R, и FreeMath. В данной работе использовался свободно распространяемый язык R.

## **1. Показатели научно-технического развития**

Для исследования были выбраны пять показателей научно-технического развития: количество патентных заявок по месту (стране) обращения; количество патентных заявок по гражданству заявителя; количество исследователей и инженеров в научной области; высокотехнологичный экспорт в долларах США по текущему курсу; количество научных публикаций. Источники: база данных Всемирного Банка (World Bank); база данных Всемирной организации интеллектуальной собственности (WIPO). Всего для анализа взяты данные по 56 странам мира из различных регионов. Длина полученных временных рядов с годичной периодичностью составила  $T=22$  шага (с 1992 по 2013 год).

## **2. Реализация метода в среде R**

Язык R, разработанный в 1993 году, распространяется под бесплатной лицензией GNU GPL 2. Существенной особенностью языка является открытая архитектура, что приводит к высокой масштабируемости, осуществляемой с помощью пакетов (библиотек, содержащих набор функций для решения специфической задачи). Официальный репозиторий The Comprehensive R Archive Network по адресу [cran.r-project.org](http://cran.r-project.org) на текущий момент включает более 7000 постоянно обновляемых и дополняемых пакетов. Язык R широко применяется в учебной и научной деятельности университетов [4].

На этапе препроцессинга решалось три задачи: интерполяция пропущенных значений ряда, экстраполяция рядов до совпадения начальных и конечных точек, нормализация рядов. Для интерполяции использовался алгоритм локально-взвешенной непараметрической регрессии `loess` [5]. Функция `loess()` включена в стандартный пакет 'stats' языка R. Для тестирования качества интерполяции составлена выборка из 10 стран, в которых были удалены участки объёмом 0.2 от общей длины ряда. Средняя относительная ошибка приближения составила 0.094.

Для экстраполяции сравнивались два алгоритма: интегрированный алгоритм авторегрессии – скользящего среднего ARIMA, и метод экспоненциального сглаживания [6].

Алгоритм, предложенный в [6] позволяет автоматизировать нетривиальную задачу подбора параметров модели ARIMA с помощью функции `auto.arima()` пакета 'forecast'. Средние относительные ошибки составили 0.123 и 0.142 (ARIMA и экспоненциальное сглаживание соответственно).

В качестве мер различия между рядами показателей использовался класс непараметрических мер, основанных на попарной разности исходных точек ряда [2]. Для кластеризации было выбрано многомерное расстояние, основанное на алгоритме динамического выравнивания временных рядов DTW [7], которое при решении данной задачи показало наибольшую эффективность [3]. Автоматический расчёт данного расстояния реализован с помощью функции `dtw()` пакета 'dtw'.

Для уточнения многомерного расстояния DTW в работе предложено обобщения одномерной меры SORT на многомерный случай [8]. Для этого автором предложен коэффициент временной корреляции двух временных рядов  $X_T, Y_T$ .

Коэффициент временной корреляции для наборов рядов  $\{X_{1T}, \dots, X_{KT}\}, \{Y_{1T}, \dots, Y_{KT}\}$  предлагается ввести как

$$\text{CORT}^K[\{X_{1T}, \dots, X_{KT}\}, \{Y_{1T}, \dots, Y_{KT}\}] = \frac{\sum_{k=1}^K \sum_{t=1}^T (X_{k_{t+1}} - X_{k_t})(Y_{k_{t+1}} - Y_{k_t})}{\sum_{k=1}^K \left[ \sqrt{\sum_{t=1}^T (X_{k_{t+1}} - X_{k_t})^2} \sqrt{\sum_{t=1}^T (Y_{k_{t+1}} - Y_{k_t})^2} \right]}$$

Для расчёта расстояний применялись встроенные объекты классов `ts` (временной ряд), и `dist` (матрица попарных расстояний). Возможность встраивать пользовательскую функцию вычисления расстояния в функцию кластеризации, а также возможности пакета `'tsclust'` [9], позволяет в значительной степени автоматизировать процесс кластеризации. Метод силуэта [10] позволяет определить оптимальное в смысле суммы внутрикластерных расстояний количество кластеров исследуемой модели. Описанные алгоритмы реализованы в виде функций и в простой форме совмещаются с результатами предыдущих этапов.

### 3. Результаты кластеризации

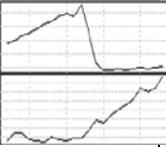
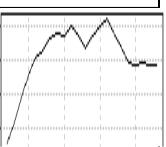
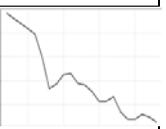
Результаты оценки эффективности кластеризационных моделей приведены в Таблице 1. Предложенная мера CORT показывает большую точность за счёт более корректной кластеризации 5 стран, расположенных на границе двух кластеров.

**Таблица 1.** Сравнение эффективности используемых мер.

Мера	Коэффициент сходства	Силуэт	Время расчёта (сек)
Многомерный DTW	0,861	0,161	<b>180,9</b>
Многомерный CORT	<b>0,874</b>	0,176	233,8

Таким образом, мера CORT эффективна в случаях, когда важно учитывать распределение приграничных элементов. В результате произведено разбиение исследуемой выборки на четыре устойчивых интерпретируемых кластера (см. Таблицу 2).

**Таблица 2.** Полученные кластеры.

#	Кол- ич.	Центр	Страны	Описание	Пример
1	26	Бразилия	США, Китай, Германия, Франция, Бразилия, Турция	Устойчивый рост по 4 или 5 показателям	
2	16	Чехия	Эстония, Латвия, Литва, Польша, Болгария	Падение по первому пока- зателю, рост по остальным	
3	4	Швеция	Финляндия, Великобрита ния, Япония	Рост сменяется стагнацией / медленным па- дением	
4	10	Парагвай	Россия, Укра- ина, Панама, Шри-Ланка	Флуктуации или падение по 4 или 5 по- казателям	

Если установить отношения доминирования, основанное на наибольшей предпочтительности устойчивого роста по всем показателям, то четвёртый кластер следует считать наихудшим. Динамика показателей России очень схожа с динамикой Украины. В частности, количество исследователей и учёных уменьшалось почти линейно на протяжении последних двадцати лет. Устойчивый рост наблюдается только в области высокотехнологического экспорта, остальные демонстрируют падение или значительную волатильность. Это может свидетельствовать о недостатке мер стимулирования научно-технического комплекса.

Решение задачи с учетом тестирования алгоритмов и визуализации реализованы с помощью примерно 600 строк кода языка R. При необходимости реализации функций самостоятельно, объём необходимой работы увеличивается примерно на

800 строк, то есть, время разработки метода при использовании среды R сокращается более чем в два раза.

## **Выводы**

Проведённое исследование позволило выделить существенные особенности динамики развития научно-технического прогресса, и создать инструменты для комплексного анализа и прогнозирования в каждой из выбранных групп. Показано, что язык R может служить эффективным инструментом для прототипирования методов и алгоритмов анализа данных с учетом специфики предметных областей. Дополнительные возможности масштабирования при использовании языка R могут быть получены при использовании технологии Oracle R Enterprise© – т.е. реализации совместных возможностей вычислительных серверов, баз данных и модулей языка R.

## **Список литературы**

1. *Геловани В.А, Бритков В.Б.* Методология глобального моделирования в анализе социально-экономических проблем // Системные исследования. Методологические проблемы: Ежегодник 2013-2014. Вып. 37. - М.: ЛЕНАНД, 2014. С. 49–63.
2. *Зайцев Р.Д.* Применение непараметрических мер различия между временными рядами для кластеризации показателей научно-технического развития // Вопросы науки: Современные технологии и технический прогресс: сборник статей по материалам IV международной научно-практической конференции. – Воронеж: Вэлборн, 2015. С. 13–19.
3. *Зайцев Р.Д.* Исследование эффективности многомерной кластеризации временных рядов для анализа динамики научно-технического развития // Перспективы развития информационных технологий: сборник статей по матери-

алам XXV Международной научно-практической конференции. – Новосибирск: ЦРНС, 2015. С. 7–12.

4. *Fox, J., Andersen, R.* Using the R statistical computing environment to teach social statistics courses // McMaster University, 2005.
5. *Cleveland W.S.* Robust Locally Weighted Regression and Smoothing Scatterplots // Journal of the American Statistical Association, 1979. 74 (368). P. 829–836.
6. *Hyndman R. J., Khandaka Y.* Automatic Time Series Forecasting: The forecast Package for R // Journal of Statistical Software, 2008. V. 27, Issue 3.
7. *Holt G., Reinders M.*, Multi-dimensional dynamic time warping for gesture recognition. // Annual Conference on the Advanced School for Computing and Imaging, 2007.
8. *Chouakria D.A., Nagabhushan P.N.* Adaptive Dissimilarity Index for Measuring Time Series Proximity // Advances in Data Analysis and Classification, 2007. 1(1). P. 5–21.
9. *Montero P., Vilar J. A.* *TSclust*: An R Package for Time Series Clustering // Journal of Statistical Software, 2014. Vol. 62, I. 1.
10. *Rousseeuw P.J.* *Silhouettes*: A graphical aid to the interpretation and validation of cluster analysis // J. Comput. Appl. Math., 1987. Vol. 20. P. 53–65.

# Описание распространения информации в сетях социальных связей на основе теории перколяции

С.А. Лесько<sup>19</sup>, Д.О. Жуков<sup>20</sup>

**Аннотация:** В работе рассматривается новая теоретическая модель описания состояний настроений и взглядов в сетях социальных связей, а также возможность практического использования полученных моделей (с привлечением данных социологических исследований) для прогнозирования поведения групп людей в обществе. В работе показано, что для достижения поставленной цели могут быть использованы методы теории перколяции (например, для определения пороговых значений уровней негативных настроений в обществе и исследования кластеризации общества на отдельные, связанные по настроениям и взглядам, группы граждан). Полученные модели показывают, что доля от 0,09 до 0,15 граждан, имеющих негативное настроение, уже является критической для возникновения социальных протестов.

**Ключевые слова:** теория перколяции, структура социальной сети, негативное влияние, кластеризация общества

## Введение

Исследования процессов социальной динамики состояний в обществе, является весьма актуальной и важной задачей. К современным подходам описания процессов в социальных сетях можно отнести, например модели на основе случайных графов социальных сетей [1]. Использование таких моделей позволяет избежать проблемы близкого вырождения (отсутствия дальнего действия по нескольким последовательным связям между членами социальной сети), которое присуще использованию моделей графа Маркова. Модели случайного графа на практике показывают более высокие уровни транзитивности

---

<sup>19</sup> МИРЭА; 1983 г.р.; к.т.н.; sergey@testor.ru

<sup>20</sup> МИРЭА; 1965 г.р.; д.т.н.; zhukovdm@yandex.ru

связей сети, которые невозможно оценить на основе параметров гомогенных моделей графа Маркова.

Модель смешанного членства в стохастически формирующихся группах рассматривается в работе [2]. Данная модель основывается на рассмотрении измерений присутствия или отсутствия связей между парой объектов. Анализ вероятностных изменений между парами объектов требуют специальных предположений, например независимости или предположения непостоянства данной связи (смешанного членства в стохастически формирующихся группах).

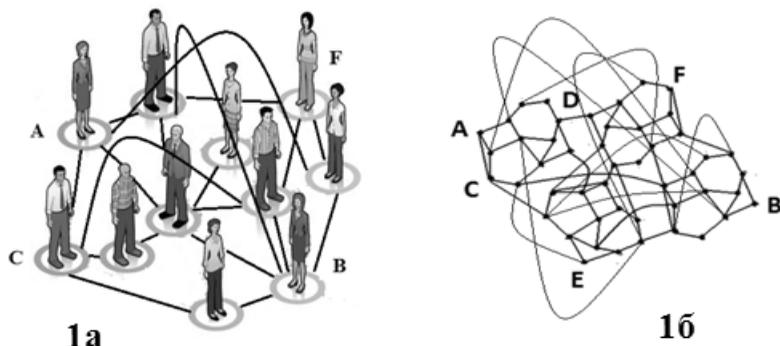
В модели соразвития выбора и влияния в социальных сетях [3] рассматривается что число узлов и топология сети (структура связей), может измениться со временем. Недостатком данной модели является то, что она явно рассматривает связи между всеми пары узлов, приводящих к квадратичной сложности при расчетах изменения численности членов в различных группах. Однако сети реального мира являются разреженными (большинство участников не имеют парных связей и вообще число их связей само по себе является случайным). Введение разреженности в модель [3], а также учет случайного характера числа связей для каждого члена сети могло бы значительно увеличить эффективность данной модели.

В модели [4] проводится анализ больших многомодальных социальных сетей. Суть разработанного в [4] подхода заключается в том, что интерактивные социальные сети часто одновременно включают многократные отношения, люди могут построить социальную сеть добавляя друг друга как друзей, они также могут сформировать несколько неявные социальных сетей (многомодальность). Авторы [4] используют методику анализа структуры развивающегося во времени, имеющего многомодальность графа социальной сети. Применение данного подхода к реальным структурам показало, что существует временная онлайн регулярность в социальных взаимодействиях людей. Кроме того, были обнаружены корреляции между возникновением дружбы между участниками и параметрами настройки интерактивной социальной сети.

Следует отметить что, ни одна из рассмотренных моделей не может считаться полностью адекватной для описания динамики социальных процессов. Все это показывает необходимость поиска и создания новых моделей описания социальной динамики и управления информацией в сетях социальных связей и отношений. По нашему мнению определенный прогресс может быть достигнут при разработке моделей использующих для описания структуры сетей случайные графы, как наиболее близко отражающие их реальное строение, и методы теории перколяции.

## 1. Структура сети социальных связей

Узлами социальной сети являются отдельные люди, а ребрами – коммуникативные связи, число которых может иметь произвольное значение. На Рис. 1а и 1б представлены пример социальных связей и их формализованная структурная схема.



**Рис. 1.** 1а. - пример социальных связей между индивидуумами в обществе, 1б - структура сети социальных связей между отдельными личностями (А,В,С,Д и т.д.) в обществе

В сети с показанной на Рис. 1б структурой, распространение информации любого типа и характера может происходить различными путями, проходя через различные узлы сети.

И таких маршрутов может одновременно существовать достаточно большое количество. Число связей для каждого человека и их среднее количество в расчете на всех граждан может быть установлено с помощью мониторинга или социологических исследований существующих настроений.

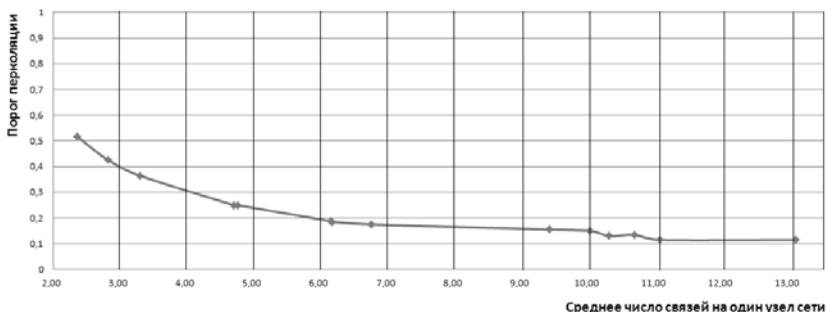
## **2. Перколяция в социальной сети**

В теории перколяции, доля проводящих (не заблокированных) узлов при которой возникает проводимость между двумя различными произвольно выбранными узлами сети называется порогом перколяции (протекания). Свободное распространение негативной информации без блокировки между двумя любыми, произвольно выбранными узлами (индивидуумами) социальной сети можно рассматривать как критерий угрозы.

Для изучения случайных сетей с множеством связей аналитических моделей описания перколяционных процессов не существует и их исследование возможно только численными методами моделирования. Для этого необходимо сначала построить структурную модель социальной сети (см. Рис. 1б). Затем, выбрать пару произвольных различных узлов и с помощью методов численного моделирования определить при какой доли не заблокированных узлов в рассматриваемой сети появляется свободный путь между узлами А и В. Затем, аналогичным образом эта процедура проводится для других произвольных пар узлов (в нашем случае пары узлов С и D, Е и F на Рис. 1б и т.д.). После этого с усреднением результатов по отдельным экспериментам, нами проводилось определение среднего значения порога перколяции по всем рассматриваемым парам узлов.

Результаты проведенного нами численного моделирования нахождения порога перколяции для случайных сетей с множеством путей между узлами и различным средним числом связей на один узел представлены на Рис. 2. Из него видно, что с ростом среднего числа связей, приходящихся на один узел сети, порог перколяции начинает монотонно стремиться к некоторому минимальному значению.

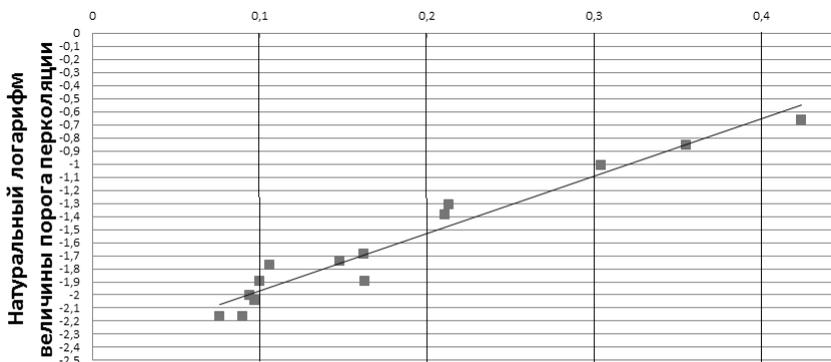
Графический вид зависимости можно описать функцией вида:  $P(x) = P_0 e^{-x}$ , где  $P(x)$  – величина порога перколяции при среднем числе связей на узел, равном некоторому значению  $x$ ,  $P_0$  – величина порога перколяции при бесконечно большом числе связей на один узел.



**Рис. 2.** Зависимость величины порога перколяции в случайной сети от среднего числа связей на один её узел

Как показывает Рис. 3, полученные данные хорошо линейризуются в координатах:  $\ln P(x)$  в зависимости от  $z=1/x$  (натуральный логарифм порога перколяции – это величина, обратная среднему числу связей  $x$ , приходящихся на один узел), что подтверждает возможность использования функцией вида:  $P(x) = P_0 e^{-x}$ .

Точки на Рис. 3 обозначают экспериментальные данные, а сплошная линия соответствует линейной зависимости:  $y = 4,39z - 2,41$ , с очень большим значением коэффициента корреляции, равным 0,95.



Величина, равная единице деленой на число связей, приходящихся на один узел

**Рис. 3.** Зависимость логарифма величины порога перколяции в случайной сети от величины обратной среднему числу связей на один её узел

При  $z=1/x=0$  (соответствует случаю  $x=\infty$ ) получаем:  $y=\ln P(0)=-2,41$ , а величина самого порога перколяции при бесконечно большом числе связей на узел  $P(0)$  будет равна 0,09. Следует отметить, что логически она должна стремиться к 0, однако полученный результат может быть объяснен тем, что при очень большом числе связей может происходить изменение закона зависимости порога перколяции от числа связей. Таким образом, для случайной сети с бесконечно большим числом связей на один узел достаточно иметь долю активированных узлов равную 0,09 от общего числа, чтобы возникла проводящая цепочка узлов и сеть решала поставленную информационную задачу. При среднем числе связей равном 100 – порог перколяции будет равен 0,094, а при 10 – 0,139.

## Выводы

1. Общество можно рассматривать как глобальную социальную сеть, имеющую произвольную случайную топологию, в которой между отдельными узлами существует произвольное случайное число связей. А протекающие в такой сети

социальные процессы можно описывать на основе аппарата теории перколяции.

2. При определенной доле негативно настроенных людей (узлов сети) может создаваться условие для беспрепятственного распространения информации между двумя любыми, произвольно выбранными узлами (индивидуумами) социальной сети, что можно рассматривать как достижение порога перколяции. При среднем числе связей равном 100 – порог перколяции будет равен 0,094, а при 10 – 0,139. Таким образом, можно считать, что доля от 0,09 до 0,15 граждан имеющих негативное настроение является уже критической для возникновения социальных потрясений.

## Список литературы

1. *Robins G., Snijders T., Wang P., Handcock M., Pattison P.* Recent developments in exponential random graph models for social networks. // *Social Networks*, 2007. Vol. 29. P. 192–215.
2. *Airoldi E. M., Blei D. M., Fienberg S. E., and Xing E. P.* Mixed membership stochastic blockmodels // *J. Mach. Learn. Res.*, 2008. Vol. 9. P. 1981–2014.
3. *Cho Y.-S., Steeg G. V., and Galstyan A.* Co-evolution of selection and influence in social networks // In W. Burgard and D. Roth, editors, *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2011*. – San Francisco, California, USA. – AAAIPress, 2011.
4. *Du N., Wang H., and Faloutsos C.* Analysis of large multi-modal social networks: patterns and a generator // In *Proceedings of the 2010 European Conference on Machine learning and knowledge discovery in databases: Part I, ECML PKDD'10*. P. 393–408.

# Структура информационного пространства: директор и самоподобие

С.А. Лесько<sup>21</sup>, Д.О. Жуков<sup>22</sup>

**Аннотация:** В представленной работе для описания структуры новостного информационного пространства вводится понятие директора: условной оси, положение которой определяется усреднением направленных векторов, задающих положение центров всех новостных кластеров. Для описания динамики достижения новостного события вводятся понятия понижающего и повышающего трендов изменения состояния информационного пространства. Анализ самоподобия в поведении директоров и трендов на основе предлагаемой модели может позволить выявить наличие периодичности в приближении или удалении состояния информационного пространства от точки прогнозируемого события.

**Ключевые слова:** информационное пространство, новостной кластер, директор информационного пространства, самоподобие процессов в информационном пространстве

## Введение

Одним из направлений работы с большими объемами текстовых данных является их анализ с целью прогнозирования появления новых информационных событий, что является очень трудной задачей. Прогнозирование появления новых информационных событий осложняется необходимостью поиска скрытых закономерностей в больших объемах слабоструктурированной гетерогенной информации и преодолением трудностей связанных с проблемой измеримости параметров протекающих информационных процессов. Априори все данные являются по своему характеру гетерогенными (*имеют разный формат представления и неоднозначные единицы измерения*). При создании модели необходим математический аппарат, который

---

<sup>21</sup> МИРЭА; 1983 г.р.; к.т.н.; sergey@testor.ru

<sup>22</sup> МИРЭА; 1965 г.р.; д.т.н.; zhukovdm@yandex.ru

позволил бы формализовать характер данных и привести их к единой шкале измерений.

Разработка методик прогнозирования событий представляет значительный интерес и в этой связи можно упомянуть работу [1], в которой представлена модель временных зависимостей в потоках событий. Авторы этой статьи вводят кусочно-постоянную модель интенсивности событий для изучения временных зависимостей в потоках событий, при этом используется Байесовский подход и распределение Пуассона к описанию выборки важности будущих событий. Это позволило разработать алгоритм, позволяющий изучать нелинейные временные зависимости для предсказания будущих событий с использованием дерева решений.

На наш взгляд, значительный прогресс в данной области может быть достигнут, если удастся установить, что процессы в информационном пространстве обладают свойствами самоподобия.

## **1. Структура и директор информационного новостного пространства, самоподобие процессов**

Используя формализованные представления данных в информационном пространстве, попробуем создать метод прогнозирования новостных событий, основанный на предположении о том, что процессы, протекающие в пространстве информационных событий обладают свойством самоподобия.

В представленной работе мы описываем разработанный нами для прогнозирования информационных событий подход, основанный на следующем. В реальном мире существуют множественные причинно-следственные связи, то при отображении событий в информационное пространство эти связи также должны хотя бы частично сохраняться. Любое событие может быть описано в информационном пространстве некоторым новостным кластером, имеющим свои собственные характери-

стики. Информационное пространство является «зеркалом» физического мира, отображающего его основные свойства и взаимосвязи событий. Информационное пространство, так же как и реальное физическое, обладает памятью и способностью к самоорганизации. Несмотря на то, что прогнозируемое новостное событие является априори неизвестным, можно вербально описать его в информационном пространстве, создавая новостной образ, а затем построить модель возможной трансформации уже существующих кластеров к заданному образу события.

В информационном пространстве можно выделить две взаимосвязанные подсистемы:

- «медленную», в которой обрабатывается информация, медленно изменяющуюся или почти не изменяющаяся с течением времени (научные, культурные, религиозные, художественные и т.д. знания);
- «быструю», к которой можно отнести информацию политического, экономического, спортивного и т.д. характера.

Каждая из подсистем содержит определенные наборы кластеров информации, со своими векторами, задающими их положение.

В рамках нашей модели мы предлагаем ввести понятие директор. Директором будем называть условное направление в информационном пространстве, которое определяется взаимным усреднением направлений всех рассматриваемых векторов. Направление директора может быть рассчитано с помощью метода наименьших квадратов для отклонения углов векторов, задающих положение центров информационных кластеров от условного направления, которое принимается за директор.

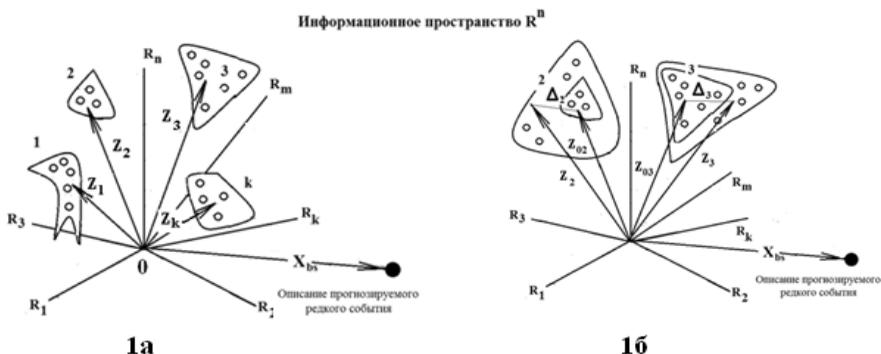
При наличии в зависимости поведения от времени величин рассматриваемых директоров самоподобия, для его описания может быть применена теория Харста [2], суть которой заключается в том, что непрерывный стохастический процесс  $X(t)$ , считается статистически самоподобным с параметром Харста (Hurst)  $H$  ( $0,5 \leq H \leq 1$ ), если для любого положительного числа  $a$ ,

случайные процессы  $X(t)$  и  $a^{-H}X(at)$  будут иметь одинаковые статистические свойства. Значение  $H=0.5$  показывает отсутствие самоподобности, а значения  $H$  близкие к единице показывают её большую степень.

## 2. Вектор прогнозируемого события и тренды процессов в информационном пространстве

Проведем в какой-то момент времени  $t$  кластеризацию информационного пространства по различным новостным событиям и определить значения величин векторов  $(z_1, z_2, z_3, \dots, z_k, z_j)$ , задающих положение центров этих кластеров в данный момент времени (см. Рис. 1). Далее проведем вербальное описание прогнозируемого новостного события и таким образом зададим его вектор  $X_{bs}$  в информационном пространстве (см. Рис. 1а).

Поскольку мы предполагаем, что в информационном пространстве уже имеются некоторые данные о предстоящем новостном событии, то должно существовать и отображение имеющихся групп новостных событий, на событие которое мы пытаемся прогнозировать. Априори мы не можем точно указать математическую формулу такого отображения, но оно должно иметь однозначный характер и сохранять причинно-следственные связи и последовательности событий. Однако при этом возможно искажение масштабов интервалов между событиями. Выберем в качестве отображения, нахождение проекций  $x_j$  векторов, задающих положение центров информационных кластеров в данный момент времени  $z_1, z_2, z_3, \dots, z_k, z_j$ , на направление вектора  $X_{bs}$ , определяющего появление прогнозируемого события. В данном случае мы предполагаем, что данные проекции являются источником формирования редкого события на оси  $X$ , задающей его появление. Каждая из проекций  $x_k$  определяется как произведение величины соответствующего вектора  $z_k$  и косинуса угла между направлениями векторов  $z_k$  и  $X_{bs}$ ; при этом  $x_k = z_k \cdot \cos(\alpha_k)$ .



**Рис. 1.** 1а. Кластеризация информационного пространства по новостным группам и определение в нем положения редкого события. 1б. Изменение положения центров кластеров за интервал времени измерения  $t_0$ .

Спустя некоторый интервал времени (назовем его интервал измерения  $t_0$ ) величины и положения векторов, задающие положения центров информационных кластеров изменяются на какие-то случайные значения  $\Delta_j$  ( $j$  – обозначает рассматриваемый вектор). На Рис. 1б в качестве примера показано изменение за время  $t_0$  для векторов  $z_2$  и  $z_3$ , задающих положение центров кластеров номер 2 и номер 3. Вектора  $z_{02}$  и  $z_{03}$  определяют положение центров новостных кластеров в информационном пространстве в момент времени  $t$ , а вектора  $z_2$  и  $z_3$  спустя интервал времени  $t_0$  (момент времени  $t + t_0$ ). В данных случаях  $\Delta_2 = z_2 - z_{02}$  и  $\Delta_3 = z_3 - z_{03}$ . Аналогичным образом определяются изменения положения центров для всех кластеров в информационном пространстве.

Величины  $x_{02}$ ,  $x_{03}$ ,  $x_2$  и  $x_3$  будут задавать значения соответствующих проекций векторов, определяющих положение центров новостных кластеров 2 и 3, на направление прогнозируемого события, в моменты времени  $t$  и  $t + t_0$ .

Следует отметить, что некоторые величины проекций  $x_k$  могут оказаться больше предыдущих значений  $x_{0k}$  (для одной и той же группы новостных данных), а некоторые меньше, т.е. в информационном пространстве одновременно сосуществуют

два тренда поведения. Один на увеличение значений проекций, другой на их уменьшение, что необходимо учесть в рамках разрабатываемой модели.

Введем для любого момента времени понятие среднего значения  $\bar{x}_i$  всех величин проекции векторов, задающих положение центров новостных кластеров в информационном пространстве на направление оси прогнозируемого события. Для группы из  $K$  новостных кластеров в момент времени  $t$  среднее значение  $\bar{x}_t$  определяется следующим образом:

$$\overline{x(t)} = \frac{\sum_{j=1}^K x(t,j)}{K}, \quad (1)$$

где  $x(t, j)$  обозначают соответствующие значения проекций векторов, задающих положение центров новостных кластеров, на направление оси прогнозируемого события, в данный момент времени  $t$ . Спустя интервал времени  $\tau_0$ :

$$\overline{x(t + \tau_0)} = \frac{\sum_{j=1}^K x(t + \tau_0, j)}{K} \quad (2)$$

Чтобы учесть тренды увеличения и уменьшения значений величин проекций векторов, задающих положение центров новостных кластеров, на направление оси прогнозируемого события можно поступить следующим образом. На основании анализа значений  $x(t, j)$  и  $x(t + \tau_0, j)$  разделим множество  $x(t, j)$  на две подгруппы, в одной  $x(t, j)I$  будут все новостные кластеры, для которых за интервал времени  $\tau_0$  произошло уменьшение значений величин проекций  $x(t + \tau_0, j)$  (обозначим число таких кластеров как  $R$ ), а во второй  $x(t, j)II$  – увеличение (обозначим число таких кластеров как  $K-R$ ), и найдем для каждой из них средние значения проекций векторов, задающих положение центров этих новостных кластеров:

$$\overline{x(t, j)}_I = \frac{\sum_{j=1}^R x(t, j)_I}{R}, \quad \overline{x(t, j)}_{II} = \frac{\sum_{j=1}^{K-R} x(t, j)_{II}}{K-R} \quad (3)$$

Далее, мы предлагаем следующий подход к учету трендов увеличения и уменьшения значений величин проекций векторов, задающих положение центров новостных кластеров, на направление оси прогнозируемого события. Поскольку при учете трендов имеет смысл говорить об усредненных величинах, то будем рассматривать переход за интервал времени  $\tau_0$  в

точку  $\overline{x(t + \tau_0)}$  из точки  $\overline{x(t, j)}_I$ , которая находится на оси прогнозирования события правее точки  $\overline{x(t + \tau_0)}$  и точки  $\overline{x(t, j)}_{II}$  которая находится левее  $\overline{x(t + \tau_0)}$ . Сами по себе переходы являются случайными событиями, а их величины можно определить следующим образом:

$$\xi(t) = \overline{x(t, j)}_I - \overline{x(t + \tau_0)}, \quad \varepsilon(t) = \overline{x(t + \tau_0)} - \overline{x(t, j)}_{II} \quad (4)$$

После следующего шага  $\tau_0$  определяем новые значения  $\xi(t + \tau_0)$  и  $\varepsilon(t + \tau_0)$ :

$$\begin{aligned} \xi(t + \tau_0) &= \overline{x(t + \tau_0, j)}_I - \overline{x(t + 2\tau_0)} \\ \varepsilon(t + \tau_0) &= \overline{x(t + 2\tau_0)} - \overline{x(t + \tau_0, j)}_{II} \end{aligned} \quad (5)$$

На любом шаге  $n$  величины  $\xi_{t+n\tau}$  и  $\varepsilon_{t+n\tau}$  могут принимать случайные (или почти случайные) значения, однако в их поведении могут наблюдаться характерных особенностей (например, зависимости  $\xi_{t+n\tau}$  и  $\varepsilon_{t+n\tau}$  от времени, могут обладать самоподобием).

Анализ самоподобия в трендах увеличения и уменьшения может позволить выявить наличие периодичности в приближении или удалении от прогнозируемого события в информационном пространстве.

## Список литературы

1. *Gunawardana A., Meek C., Xu P.* A Model for Temporal Dependencies in Event Streams // *Advances in Neural Information Processing Systems*, 2011. P. 1962–1970.
2. *Clegg R.G.* A practical guide to measuring the hurst parameter. // *Computing science technical report.* – 2005. – № CS–TR–916. P. 125–138.

# Концептуальная модель коллективов интеллектуальных агентов для решения сложных задач

С.В. Листопад<sup>23</sup>

**Аннотация:** Существующие интеллектуальные системы поддержки принятия решений строятся преимущественно в рамках парадигмы компьютерного моделирования процессов и явлений в памяти одного человека, что обуславливает сложность их применения для решения неоднородных задач. Предлагается подход к построению коллективов агентов, поддерживающих принятие решений, в рамках синергетической парадигмы искусственного интеллекта. Разработанные к настоящему моменту гибридные интеллектуальные системы, реализующие отдельные элементы предлагаемого подхода, демонстрируют релевантность сложным практическим задачам, а также перспективность данного направления исследований.

**Ключевые слова:** коллектив интеллектуальных агентов; система поддержки принятия решений; гибридная интеллектуальная система

## Введение

В настоящее время компьютерные технологии по-прежнему остаются средой реализации методов, а не инструментальным средством их синтеза. В результате, на практике со сложными задачами и уникальными ситуациями справляются только коллективы экспертов под руководством лица, принимающего решения. Взаимодействуя в ходе обсуждения, эксперты обмениваются данными, знаниями, объяснениями и частичными решениями общей задачи. При этом возникают процессы самоорганизации, направляемые отношениями кооперации, компромисса, содействия, конкуренции, приспособления и др., и определяющие возможность возникновения синергетического

---

<sup>23</sup> ФИЦ ИУ РАН (Калининградский филиал); 1984 г.р.; к.т.н.; ser-list-post@yandex.ru

эффекта, когда коллективное, интегрированное решение оказывается лучше частных мнений экспертов. Аналогично экспертным системам, позволившим компьютеру рассуждать «не хуже» одного человека, актуальны интеллектуальные системы, работающие в условиях сложных задач не хуже коллектива специалистов.

## 1. Понятие коллектива

Коллектив можно определить как организованную систему субъектов, в которой каждый из них обладает собственными ценностями, целями и задачами, согласованные и подчиненные общим для системы ценностям, целям и задачам [1, 2]. Каждый член коллектива обладает индивидуальной функцией полезности, которую он стремится максимизировать при максимизации глобальной функции полезности, характеризующей работу всего коллектива [3].

Может быть выделено девять основных атрибутов коллектива:

1. Предназначение, определяющее цели и задачи коллектива.

2. Наличие лидера, который назначает задачи, управляет, координирует и мотивирует экспертов, а также планирует, организует и оценивает работу коллектива.

3. Взаимодействие – процесс обмена данными, информацией, знаниями, объяснениями, частичными решениями между членами коллектива, а также с внешней средой коллектива, обеспечивающий возникновение ряда макроскопических эффектов [4].

4. Принятие решений – процесс выбора между двумя или более альтернативами при решении возникающих проблем.

5. Культура – нормы и ценности, принятые членами коллектива, способы их взаимодействия друг с другом.

6. Коллективная мотивация – психологически членство в успешном коллективе может быть источником мотивации и ресурсом для решения задач.

7. Конфликт – напряжение, испытываемое отдельными лицами или группами внутри коллектива в результате воспринимаемых различий между ними, причем конфликты относительно задач обычно повышают эффективность работы коллектива, а персонализированные конфликты – снижают.

8. Совещания, усиливающие интенсивность процессов взаимодействия и принятия решений.

9. Самоконтроль – персональные навыки, которые требуются каждому участнику, чтобы выполнить свои обязательства перед коллективом и успешно работать с другими членами коллектива.

Один из первых шагов по построению интеллектуальные системы как коллектива агентов, моделирующего работу реальных коллективов экспертов, – разработка его концептуальной модели.

## 2. Модель коллектива интеллектуальных агентов

На основе перечисленных особенностей коллектива построим концептуальную модель коллектива интеллектуальных агентов (КИА), описывающую его основные элементы, которые требуется реализовать в виде программного комплекса. Модель – расширение концептуальной модели системы поддержки принятия решений [4], рассматривающее КИА на двух уровнях. На микроуровне описываются взаимосвязи и функциональные возможности членов КИА, на макроуровне – поведение и характеристики КИА в целом. Концептуальные модели ниже записываются в аксиоматической теории схем ролевых концептуальных моделей [5].

Микроуровневая концептуальная модель КИА – выражение:

$$\begin{aligned} \square_{\text{iht}} = & R^{\text{res res}}(PRT, \text{env}) \circ R^{\text{res res}}(PRT, PRT) \circ R^{\text{res res}}(PRT, \text{ont}) \circ \\ & \circ R^{\text{res res}}(\text{ont}, PRT), \quad PRT \subseteq RES, \quad \text{env} \in RES, \quad \text{ont} \in RES, \end{aligned} \quad (1)$$

где  $PRT = \{prt_1, \dots, prt_n, prt^{dm}\}$ ,  $PRT$  – множество агентов-участников КИА, включающее лидера, руководителя коллектива –

агента, принимающее решения  $prt^{dm}$ ;  $env$  – внешняя среда;  $ont$  – модель предметной области КИА;  $R^{res\ res}$  – множество отношений «ресурс – ресурс» участников КИА между собой и с внешней средой, а также между агентами-участниками КИА и моделью его предметной области;  $RES$  – множество концептов типа «ресурс».

Каждый агент-участник  $prt_i \in PRT$  КИА имеет собственную функцию полезности  $pr^{uf}$ , преследует цель  $pr^{ssu}$ , которая может совпадать или вступать в противоречие с целями других агентов и коллектива в целом, решает  $act^{sol}$  подмножество задач  $PRB_i \subseteq PRB$  множества задач  $PRB$ , решаемых коллективом [4]. Взаимодействуя при решении задач, агенты обмениваются данными  $pr^{dat}$ , знаниями  $pr^{knw}$ , объяснениями  $pr^{exp}$  и частичными решениями  $pr^{dec}$  общей задачи, т.е. выполняют множество действий по передаче  $ACT^{itr}$  и получению  $ACT^{iac}$  информации, множество функций  $ACT^{prt}$ , оказывают давление на других агентов КИА и окружающую среду, выполняя действия  $ACT^{conf}$ . Каждый агент-эксперт обладает собственной моделью  $res^{mod}$  предметной области, а также собственным набором методов  $RES^{met}$  решения задач. Агенты-участники успешного КИА должны обладать навыками самоконтроля  $act^{sfc}$ , организованности  $act^{org}$ , терпения  $act^{pat}$ , управления стрессом  $act^{stm}$  и уверенности  $act^{efd}$ . Для успешного решения сложных задач в состав КИА должны входить агенты, моделирующие экспертов различных специальностей, с разными наборами методов решения задач, т.е.  $RES_i^{met} \neq RES_j^{met}$ , где  $i, j = 1, \dots, n, i \neq j$  – номер агента во множестве  $PRT$ . Модель агента-участника записывается выражением:

$$\begin{aligned}
prt_i = & r_1^{res\ pr} (prt, pr^{df}) \circ r_1^{res\ pr} (prt, pr^{gsu}) \circ r_2^{res\ act} (prt, act^{sol}) \circ \\
& \circ r_1^{act\ res} (act^{sol}, PRB_i) \circ r_1^{res\ pr} (prt, pr^{dat}) \circ r_1^{res\ pr} (prt, pr^{knw}) \circ \\
& \circ r_1^{res\ pr} (prt, pr^{exp}) \circ r_1^{res\ pr} (prt, pr^{dec}) \circ r_2^{res\ act} (prt, ACT^{prt}) \circ \\
& \circ r_2^{res\ act} (prt, ACT^{irr}) \circ r_2^{act\ pr} (ACT^{irr}, pr^{dat}) \circ r_2^{act\ pr} (ACT^{irr}, pr^{knw}) \circ \\
& \circ r_2^{act\ pr} (ACT^{irr}, pr^{exp}) \circ r_2^{act\ pr} (ACT^{irr}, pr^{dec}) \circ r_2^{res\ act} (prt, ACT^{iac}) \circ \\
& \circ r_2^{act\ pr} (ACT^{iac}, pr^{dat}) \circ r_2^{act\ pr} (ACT^{iac}, pr^{knw}) \circ r_2^{act\ pr} (ACT^{iac}, pr^{exp}) \circ \\
& \circ r_2^{act\ pr} (ACT^{iac}, pr^{dec}) \circ r_2^{res\ act} (prt, ACT^{conf}) \circ r_3^{res\ res} (prt, res^{mod}) \circ \\
& \circ r_3^{res\ res} (prt, RES^{met}) \circ r_2^{res\ act} (prt, act^{sf}) \circ r_2^{res\ act} (prt, act^{org}) \circ \\
& \circ r_2^{res\ act} (prt, act^{pal}) \circ r_2^{res\ act} (prt, act^{sm}) \circ r_2^{res\ act} (prt, act^{fld}),
\end{aligned} \tag{2}$$

где  $r_1^{res\ pr}$  — отношение «иметь свойство», устанавливающие соответствие между агентом коллектива и его свойствами;  $r_2^{res\ act}$  — отношение «выполнять», связывающее субъект и действие, которое он выполняет;  $r_1^{act\ res}$  — отношение «иметь объектом», связывающее действие и объект, на который оно направлено;  $r_2^{act\ pr}$  — отношение «иметь свойство», связывающее действие с его свойством;  $r_3^{res\ res}$  — отношение «включает», связывающее целое и его части.

Следует особо рассмотреть агента, принимающего решения  $prt^{dm}$ . Модель агента, принимающего решения, записывается в следующем виде:

$$\begin{aligned}
prt_{dm} = & prt_i \circ r_2^{res\ act} (prt_{dm}, act^{pd}) \circ r_1^{act\ res} (act^{pd}, prb^u) \circ \\
& \circ r_3^{res\ res} (prb^u, PRB^h) \circ r_2^{res\ act} (prt_{dm}, act^{pas}) \circ \\
& \circ r_1^{act\ res} (act^{pas}, PRT) \circ r_2^{res\ act} (prt_{dm}, act^{acc}) \circ r_1^{act\ res} (act^{acc}, dec^u) \circ \\
& \circ r_2^{res\ act} (prt_{dm}, act^{man}) \circ r_1^{act\ res} (act^{man}, \overline{aht}) \circ r_2^{res\ act} (prt_{dm}, act^{coop}) \circ \\
& \circ r_1^{act\ res} (act^{coop}, PRT) \circ r_2^{res\ act} (prt_{dm}, act^{mot}) \circ r_1^{act\ res} (act^{mot}, PRT) \circ \\
& \circ r_2^{res\ act} (prt_{dm}, act^{pl}) \circ r_1^{act\ res} (act^{pl}, \overline{aht}) \circ r_2^{res\ act} (prt_{dm}, act^{org}) \circ \\
& \circ r_1^{act\ res} (act^{org}, \overline{aht}) \circ r_2^{res\ act} (prt_{dm}, act^{ev}) \circ r_1^{act\ res} (act^{ev}, \overline{aht}).
\end{aligned} \tag{3}$$

Дополнительно к функциям, которые выполняют рядовые участники КИА, он выполняет декомпозицию  $act^{pd}$  сложных задач  $prb^u$  на однородные подзадачи  $PRB^h$ , если это необходимо, назначает подзадачи участникам  $act^{pas}$ , принимает  $act^{acc}$  окончательное решение  $dec^u$  задачи  $prb^u$ , управляет коллективом  $act^{man}$ , координирует  $act^{coop}$  и мотивирует деятельность участников  $act^{mot}$ , планирует  $act^{pl}$ , организует  $act^{org}$  и оценивает  $act^{ev}$  эффективность КИА.

Макроуровневая модель КИА представляется выражением:

$$\square_{iht} = (PRT, env, INT, \square_{AHT}, EFF, GS) \quad (4)$$

где  $PRT$  – множество агентов КИА, описываемых концептуальной моделью (2);  $env$  – среда, в которой находится КИА;  $INT$  – множество элементов для структурирования взаимодействий между агентами;  $\square_{AHT}$  – множество моделей (1) КИА на микроуровне, соответствующих конкретным функциям агентов в КИА и отношениям между ними;  $EFF$  – множество моделей макроуровневых эффектов в КИА;  $GS$  – множество элементов целеполагания КИА.

Множество  $INT$  элементов для структурирования взаимодействий между агентами из (4) может быть представлено выражением:

$$INT = \{PROT, LANG, ont, CUL, CON, MT, DM\} \quad (5)$$

где  $PROT$  – множество используемых в КИА протоколов взаимодействия;  $LANG$  – множество языков общения принятых в КИА;  $ont$  – общая для КИА модель предметной области;  $CUL$  – множество элементов культуры, таких как принятые КИА ценности и нормы поведения агентов;  $CON$  – множество элементов для организации и разрешения конфликтов;  $MT$  – множество элементов, регламентирующих организацию совещаний;  $DM$  – множество элементов, регламентирующих процессы принятия решений.

Множество  $EFF$  концептуальных моделей макроуровневых эффектов из (4) определяется следующим образом [4]:

$$EFF = \{ad, bo, wa, ho, ge, gt, fa, re, sl, so, se, sf\} \quad (6)$$

где  $ad$  – адаптация;  $bo$  – бумеранг;  $wa$  – волна;  $ho$  – го меостаз;  $ge$  – групповой эгоизм;  $gt$  – группинк;  $fa$  – мода;  $re$  – эффект Рингельмана;  $sl$  – самообучение;  $so$  – самоорганизация;  $se$  – синергия;  $sf$  – социальная фасилитация.

Элементы целеполагания из (4) – множество:

$$GS = \{dest, uf, GL, PRB, FGL, FPRB\} \quad (7)$$

где  $dest$  – предназначение КИА;  $uf$  – функция полезности КИА;  $GL$  – множество целей КИА, которых требуется достичь для исполнения предназначения  $dest$ ;  $PRB$  – множество задач КИА, которых требуется решить для достижения целей  $GL$ ;  $FGL$  – множество функций КИА постановки целей  $GL$  на основе предназначения  $dest$ ;  $FPRB$  – множество функций КИА постановки задач  $PRB$  для достижения целей  $GL$ .

Применение предложенной концептуальной модели КИА (1) – (7) позволит направлять деятельность разработчиков КИА, релевантно подбирать из уже существующих или конструировать новые методы моделирования различных его элементов и их взаимосвязей.

### 3. Предварительные результаты

Отдельные элементы концептуальной модели КИА были реализованы при разработке гибридной интеллектуальной многоагентной системы (ГиИМАС) для решения сложной транспортно-логистической задачи [4]. Идея системы состоит в сотрудничестве технологии гибридных интеллектуальных систем, релевантных неоднородности сложной транспортно-логистической задачи, а также многоагентного подхода для моделирования в КИА эффекта самоорганизации  $so$  на основе анализа согласованности целей  $PR^{gsu}$  экспертов  $PRT$  в ГиИМАС. Модель самоорганизации ГиИМАС [4] основана на понятии нечеткой цели  $pr^{gsu}$  агента-эксперта  $prt_i$  и критерии эффективности  $uf$  всей системы. Вычисляя соотношения нечетких целей агентов, агент, принимающий решения,  $prt_{dm}$  идентифицирует текущую архитектуру ГиИМАС  $\bar{a}_{ht}$ . Основываясь на своем опыте и параметрах решаемой задачи, он может изменить цели  $pr^{gsu}$  агентов и, соответственно, архитектуру ГиИМАС  $\bar{a}_{ht}$ , чтобы повысить эффективность решения задачи. Таким образом, система динамически перестраивает алгоритм функционирования, каждый раз при работе над сложной задачей вырабатывая релевантный ей метод решения.

Анализ технико-эксплуатационных параметров показал, что в условиях лабораторного эксперимента время решения от двух до 30 мин, а повышение качества решений по сравнению с принятыми на практике подтверждено экспертами [4]. По итогам практического использования ГиИМАС на двух объектах средняя суммарная себестоимость и длительность доставки грузов в день сократилась на 7,2% и на 12,13%, среднее время построения маршрутов в день уменьшилось на 23,14%.

## **Выводы**

В работе представлена концептуальная модель КИА для решения сложных задач. Ее применение позволит направлять деятельность разработчиков КИА, обеспечивая выбор и построение релевантных методов моделирования его элементов и их взаимосвязей. Результаты лабораторных экспериментов с системой, реализующей отдельные элементы данной модели, показали эффективность их использования при решении сложных задач и перспективность данного направления исследований.

## **Список литературы**

1. *Платонов К.К.* Общие проблемы теории групп и коллективов // Коллектив и личность. – М., 1975. – С. 3 – 17.
2. *Петровский А.В., Шпалинский В.В.* Социальная психология коллектива. — М., 1978.
3. *Tumer K., Wolpert D.* A Survey of Collectives // Collectives and the design of complex systems. – New York: Springer-Verlag, 2004. – pp. 1 – 40.
4. *Колесников А.В., Кириков И.А., Листопад С.В.* Гибридные интеллектуальные системы с самоорганизацией: координация, согласованность, спор. – М.: ИПИ РАН, 2014.
5. *Колесников А.В.* Гибридные интеллектуальные системы. Теория и технология разработки. – СПб.: СПбГТУ, 2001.

# Балансировка потоков в загруженных транспортных сетях на основе стохастической модели

М.В. Неменко<sup>24</sup>, С.А. Лесько<sup>25</sup>, Д.О. Жуков<sup>26</sup>

**Аннотация:** В работе описана разработанная авторами модель управления стохастическими потоками с недетерминированными характеристиками распределения статистических параметров в городских транспортных сетях. Модель описывает зависимость вероятности блокирования отдельных узлов от характеристик дорожного движения с течением времени. Проведенное моделирование длин образующихся очередей на светофорах показывает двукратное снижение числа пробок при использовании разработанной модели «регулируемых» светофоров по сравнению с жёстко заданными режимами переключения («классическое движение»).

**Ключевые слова:** транспортная сеть, стохастическая динамика блокировки узлов транспортной сети, балансировка потоков

## Введение

По экспертным оценкам, только за один год из-за пробок, Англия теряет до 20 млрд. фунтов стерлингов, а 40 % выбросов вредных веществ в США даёт автомобильный транспорт. Одним из вариантов решения проблемы сжигания избыточного количества топлива в автомобильных пробках является создание транспортно – информационных систем, основанных на адекватных моделях и алгоритмах управления. Существующие сегодня модели динамики транспортных потоков могут быть разделена по их свойствам и типу следующим образом:

---

<sup>24</sup> МИРЭА; 1989 г.р.; nemomasha@mail.ru

<sup>25</sup> МИРЭА; 1983 г.р.; к.т.н.; sergey@testor.ru

<sup>26</sup> МИРЭА; 1965 г.р.; д.т.н.; zhukovdm@yandex.ru

1. Макроскопические (гидродинамические) модели [1] (описывают движение объектов в усредненных терминах, например таких как плотность, средняя скорость автомобиля и т.д.).

2. Кинетические (газодинамические) модели [2] (описание транспортного потока приравнивается к движению специфической жидкости, по этой причине этот класс моделей называют гидродинамическими).

3. Микроскопические модели, описывающие движение каждого автомобиля. Одной из наиболее эффективных микромоделей является модель клеточных автоматов (Cellular Automata, CA) [3].

Следует отметить, что перечисленные модели не являются универсальными и имеют ряд недостатков.

## **1. Стохастическая модель распределения потоков с недетерминированными характеристиками в городских транспортных сетях**

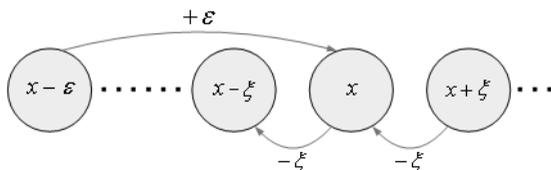
Суть разработанной нами модели состоит в следующем: если рассматривать изменение потоков машин, как случайный процесс и для каждого направления, каждого узла транспортной сети (перекрестка) задано критически допустимое число машин в очереди  $L_{i,j}$  то, можно определить вероятность  $P(L_{i,j}, t)$  того, что к моменту времени  $t$  число машин в очереди не превысит  $L_{i,j}$  (пробка не образуется).

Пусть за некоторый интервал времени  $\tau$  на перекресток  $j$ , в направлении  $i$  в очередь поступает  $\varepsilon$  машин и уезжает  $\xi$  машин. Весь процесс обработки будет складываться из отдельных шагов  $h$  имеющих продолжительность  $\tau$ , причём  $\frac{\varepsilon}{\tau} = \lambda$  – интенсивность входного потока, а  $\frac{\xi}{\tau} = \mu$  – интенсивность выходного потока машин.

Обозначим через,  $P_{x^{\varepsilon}, h}$  – вероятность того, что в оче-

реди после  $h$  шагов работы находится  $(x-\varepsilon)$  машин, а  $P_{x,h}$  – вероятность того что находятся  $x$  машин и  $P_{x+\xi,h}$  – вероятность того, что находится  $(x+\xi)$  машин. Тогда вероятность  $P_{x,h+1}$  (см. Рис. 1) того, что на  $h+1$  шаге будет находится  $x$  машин будет равна:  $P_{x,h+1} = P_{x-\varepsilon,h} + P_{x+\xi,h} \cdot P_{x,h}$ . Введем  $t=h\tau$ , где  $t$  – общее время процесса обработки и получим:

$$P(x, t+\tau) = P(x-\varepsilon, t) + P(x+\xi, t) \cdot P(x, t).$$



**Рис. 1.** Схема возможных переходов между состояниями, характеризующими число машин на  $j$ -м перекрестке, в  $i$ -м направлении на  $h+1$  шаге работы светофора

Разложим полученное уравнение в ряд Тейлора и учитывая в левой части члены, содержащие не более чем первую производную по  $t$ , а в правой не более чем вторую производную по  $x$ , получим:

$$\frac{\partial P(x,t)}{\partial t} = \frac{\lambda^2 + \mu^2}{2\mu} \frac{\partial^2 P(x,t)}{\partial x^2} - (\lambda - \mu) \frac{\partial P(x,t)}{\partial x} \quad (1)$$

Вторую производную по  $t$  можно исключить, поскольку по своему смыслу она описывает процесс, при котором сами машины могли бы быть источниками дополнительных машин. Поскольку функция  $P(x,t)$  является непрерывной можно перейти от вероятности  $P(x,t)$  к плотности вероятности  $\rho(x,t)$ , что позволит сформулировать и решить краевую задачу. При числе машин  $x=L$  в очереди на  $j$  – перекресток, в  $i$  – направлении, где  $L$  – некоторое критическое число, мы считаем, что узел обработки ( $j$  – перекресток, в  $i$  – направлении) становится перегруженным (образуется пробка). Сама вероятность обнаружить такое состояние будет отлична от 0, а плотность вероятности, определяющая поток машин в состоянии  $x=L$  необходимо положить равной 0 (мы стремимся избежать этого состояния), т.е.:  $\rho(x,t)|_{x=L} = 0$ . Второе граничное условие выбираем исходя из того,

что состояние  $x=0$  определяет простой в обработке. Сама вероятность обнаружить такое состояние будет отлична от 0, однако плотность вероятности, определяющая поток машин в состоянии  $x=0$  необходимо положить равной 0 (мы также должны стремиться избежать это состояние, т.к. оно соответствует случаю, когда светофор не закрывает данное направление, а это противоречит логике его работы), т.е.:  $\rho(x, t)_{x=0}=0$ .

Поскольку в момент времени  $t=0$  (начало расчета) на обработке может находиться  $x_0$  машин, то начальное условие зададим в виде:  $\rho(x, t = 0) = \delta(x - x_0) = \begin{cases} 1, x = x_0 \\ 0, x \neq x_0 \end{cases}$ .

Используя методы операционного исчисления для вероятности  $P(L_{i,j}, x_0/t)$  того, что к моменту времени  $t$  пробка не образуется (число машин в очереди не превысит  $L_{i,j}$ ) можно получить выражение (2):

$$P(L_{i,j}, x_0|t) = 2e^{-\frac{2b_{i,j}x_0 + b_{i,j}^2 t}{4a_{i,j}}} \sum_{n=1}^M (-1)^{n+1} \frac{e^{\frac{b_{i,j}L_{i,j}}{2a_{i,j}}} \sin\left(\frac{\pi n x_0}{L_{i,j}}\right) + \sin\left(\frac{\pi n L_{i,j} - x_0}{L_{i,j}}\right)}{\pi n + \frac{b_{i,j}^2 L_{i,j}^2}{4\pi n a_{i,j}^2}} e^{-\frac{\pi^2 n^2 a_{i,j} t}{L_{i,j}^2}} \quad (2)$$

где  $a_{i,j} = \frac{\mu_{i,j}^2 + \lambda_{i,j}^2}{2\lambda_{i,j}}$  и  $b_{i,j} = \lambda_{i,j} - \mu_{i,j}$ ,  $\mu_{i,j}$  – число машин выходящих из  $j$ -узла транспортной сети (перекресток/светофор) в  $i$ -направлении за единицу времени (выходной поток),  $\lambda_{i,j}$  – число машин входящих на узел за единицу времени (входной поток),  $t$  – время,  $x_0$  – число машин в очереди в момент начала шага работы светофора.

Решение уравнения (2) относительно времени  $t$  позволяет определить оптимальные интервалы времени включения светофоров. Однако для этого необходимо синхронизировать входящие и выходящие потоки машин на соседних перекрестках (см. уравнение (3)).

$$x_0^k{}_{i,j} = x_0^{k-1}{}_{i,j} + \frac{1}{r} \sum_{i=1}^r (\mu_{i,j}^{k-1} \tau_{i,j}^{k-1} + \Delta \lambda_{i,j}^{k-1} T_{i,j}^{k-1}) - \mu_{i,j}^k \tau_{i,j}^k \quad (3)$$

$$\lambda_{i,j}^k = \frac{x_{0,i,j}^k V_{D1}}{l_{i,j}}$$

где  $x_{0,i,j}^{k-1}$  – число машин оставшихся не пропущенными на данном направлении  $i$ , данного  $j$  - перекрестка, после выполнения  $(k-1)$  шага,  $r$  – число входящих на перекресток направлений,  $\mu_{i,j}^{k-1}$  - потоки, выходящие на  $(k-1)$  шаге по каждому из  $r$  направлений на выбранный перекресток. Любая из машина, из входящих на  $(k-1)$  шаге потоков, может равновероятно выбрать на следующем шаге  $k$  одно из направлений  $r$ , поэтому перед знаком суммы стоит числовой коэффициент  $\frac{1}{r}$ .  $T_{i,j}^{k-1}$  - время, в течение которого выбранное направление было закрыто светофором (не время открытия, а время “цикла простоя”) между двумя последовательными открытиями. Заметим, что открытие всех направлений на выбранном узле может происходить не в строго периодической последовательности. Порядок работы направлений может изменяться в зависимости от характера движения. Интервал времени между двумя последовательными открытиями одного и того же выбранного направления будет являться “циклом простоя”, величина которого  $T_{i,j}^{k-1}$  может динамически изменяться.  $\Delta \lambda_{i,j}^k$  – изменение входящего в выбранном направлении на выбранный узел потока машин за время  $T_{i,j}^{k-1}$ . Общее число машин в сети в любой момент времени суток соответствует функции числа машин от времени суток.  $\tau_{i,j}^{k-1}$  – время, в течение которого на  $(k-1)$  шаге были открыты входящие направления, пока выбранное исходящие направление было закрыто в течение времени  $T_{i,j}^{k-1}$ .  $\mu_{i,j}^k$  – поток исходящий по выбранному направлению на шаге  $k$ ,  $t_{i,j}^k$  – интервал времени включения светофоров на шаге  $k$  выбранного направления (величину которого необходимо определить при решении уравнения для определения вероятность  $P(L_{i,j}, x_0 | t)$  того, что  $k$  моменту времени  $t$  число машин в очереди не превысит  $L_{i,j}$  (пробка не образуется)).  $V_{D1}$  - рекомендуемая скорость.

## 2. Моделирование балансировки потоков в городских транспортных сетях на основе стохастической модели

На основе уравнений (1) и (2) и данных о суточном распределении числа машин на улицах г. Москвы был разработан алгоритм управления транспортной сетью и проведена проверка предлагаемого подхода. В качестве критерия эффективности был выбран показатель общей длины всех очередей на всех перекрёстках загруженной карты участка дорожной сети и смоделировано суточное движение автотранспорта. Пробкой считалась очередь автомобилей на светофоре, ожидающих разрешающего сигнала превышающая некоторое наперед заданное значение.

Проведенное исследование длин очередей (см. Рис. 2) показывает примерно двукратное снижение числа пробок при использовании модели «регулируемых» светофоров по сравнению с моделью «классическое движение» (светофоры с фиксированными фазами работы).



**Рис. 2.** Сравнение эффективности предлагаемой модели с моделью фиксированных времен переключения светофоров

### Выводы

1. Разработанные модели управления стохастическими потоками с недетерминированными характеристиками

распределения статистических параметров в городских транспортных сетях, позволяют описывать зависимость вероятности блокирования отдельных узлов с течением времени.

2. В разработанных математических моделях описаны правила обслуживания перекрестков (время переключения светофоров), учтен материальный баланс числа машин в системе и связи их потоков между соседними перекрестками, что позволяет, используя реальную карту транспортной сети создать её динамическую модель и эмулировать работу.

3. Моделирование транспортной сети города и дорожной ситуации с «управляемыми», согласно предлагаемой модели временами переключения, светофоров и с жёстко заданными режимами переключения («классическое движение») показывает двукратное снижение числа пробок при использовании разработанной модели.

## Список литературы

1. *Kerner B. S.* Introduction to Modern Traffic Flow Theory and Control. Berlin: Springer, 2009.
2. *Göttlich S., Klar A.* Model hierarchies and optimization for dynamic flows on networks. Modeling and optimization of flows on networks. Cetaro (CS), June 15–19, 2009. C.I.M.E. Courses, 2009.
3. *Maerivoet S., De Moor B.* Cellular automata models of road traffic // Physics Reports 2005. V. 419. № 1. P.1–64.

# Стохастическая динамика переходов между состояниями в сетях социальных связей, самоорганизация и скачки настроений

М.В. Неменко<sup>27</sup>, С.А. Лесько<sup>28</sup>, Д.О. Жуков<sup>29</sup>

**Аннотация:** В работе описана разработанная стохастическая модель динамики изменения состояний в сетях социальных связей и достижения порога их перехода в состояние с преобладанием негативных взглядов. Модель показывает возможность роста вероятности перехода через такой порог практически сразу после начала процесса изменения настроений в социальной сети, что связано с наличием памяти о предыдущих состояниях системы и возможности описания её самоорганизации вследствие учета в дифференциальной модели второй производной по времени. Данная модель показывает возможность скачкообразных изменений для вероятности перехода через порог негативных настроений, и учитывает наличие в её поведении волнообразных явлений. Все это очень хорошо согласуется с реальным поведением общества в ходе «цветных революций».

**Ключевые слова:** состояния социальной сети, стохастические динамические модели, самоорганизация, случайные процессы с памятью

## Введение

В настоящее время, для описания процессов социальной динамики и распространения информации используются в основном модели, пришедшие в теорию коммуникаций из биологии и экономики (экспоненциальная модель, модели Грилихеса, Коулмена, Блэкмана, Флойда, Басса, Гомпертца и многие другие).

В ряде наиболее простых случаев данные модели дают

---

<sup>27</sup> МИРЭА; 1989 г.р.; nemomasha@mail.ru

<sup>28</sup> МИРЭА; 1983 г.р.; к.т.н.; sergey@testor.ru

<sup>29</sup> МИРЭА; 1965 г.р.; д.т.н.; zhukovdm@yandex.ru

хорошие результаты, но в основном никак не соответствуют динамике процессов, которая заключается в лавинообразном нарастании количества последователей новой идеи и постепенном достижении порогового значения. Из всех упомянутых моделей наиболее соответствует реальности и имеющимся наблюдениям модель Блэкмана, дающая классическую логистическую кривую изменения во времени числа последователей данной идеи:

$$n(t) = N \frac{C \cdot e^{kt}}{1 + C \cdot e^{kt}}, \quad (1)$$

где  $N$  – общее количество потенциальных последователей идеи (весь объем общества), а  $k$  и  $C$  – константы.

По нашему мнению, описание процессов в социальных структурах возможно на основе использования стохастических феноменологических моделей, основанных на учете наличия в системе памяти и самоорганизации.

## 1. Модель стохастической динамики переходов между состояниями в сетях социальных связей

В разрабатываемом нами модели мы предлагаем описать общество, как систему, состояния которой в любой момент времени могут быть описаны некоторыми параметрами, принимающими непрерывные случайные значения. Все множество состояний будем обозначать  $X$ . Состояние, наблюдаемое в момент времени  $t$  можно обозначить, как  $x_i$  ( $x_i \in X$ ). Кроме того, введем интервал времени  $\tau_0$ , за который возможно изменение состояния  $x_i$ . В данном случае любое значение текущего времени  $t = h \tau_0$ , где  $h$  – номер шага перехода между состояниями (процесс перехода между состояниями становится квазинепрерывным с бесконечно малым временным интервалом  $\tau_0$ ),  $h=0,1,2,3, \dots, N$ . Текущее состояние  $x_i$  на шаге  $h$ , после перехода на шаге  $h+1$  может за счет случайно возникающих факторов увеличиваться на некоторую величину  $\varepsilon$ , или уменьшаться на величину  $\xi$ , и соответственно оказаться равным  $x_i + \varepsilon$ , или  $x_i - \xi$ . Введем понятие вероятности нахождения системы в том или ином состоянии.

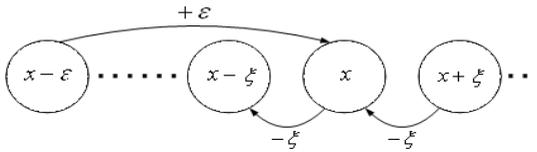
Пусть, после некоторого числа шагов  $h$  про описываемую систему можно сказать, что:

- $P(x-\varepsilon, h)$  – вероятность того, что она находится в состоянии  $(x-\varepsilon)$ ;
- $P(x, h)$  – вероятность того, что она находится в состоянии  $x$ ;
- $P(x+\xi, h)$  – вероятность того, что она находится в состоянии  $(x+\xi)$ .

После каждого шага, состояние  $x_i$  (далее индекс  $i$  можно опустить), может изменяться на  $\varepsilon$  или  $\xi$ .

Вероятность  $P(x, h+1)$  – того, что на следующем  $(h+1)$  шаге система (или процесс) окажется в состоянии  $x$  будет равна (см. Рис. 1):

$$P(x, h + 1) = P(x - \varepsilon, h) + P(x + \xi, h) - P(x, h) \quad (2)$$



**Рис. 1.** Схема возможных переходов между состояниями системы (или процесса) на  $h+1$  шаге

Поясним выражение (2) и представленную на рис.1 схему. Вероятность переход в состояние  $x$  на шаге  $h$   $P(x, h+1)$  определяется суммой вероятностей переходов в это состояние из состояний  $(x-\varepsilon) \cdot P(x-\varepsilon, h)$  и  $(x+\xi) \cdot P(x+\xi, h)$  в которых находилась система на шаге  $h$  за вычетом вероятности перехода  $(P(x, h))$  системы из состояния  $x$  (в котором она находилась на шаге  $h$ ) в другое состояние на  $h+1$  шаге.

В данном случае мы рассматриваем марковский процесс, в котором система не обладает памятью состояний. Для учета этого определим вероятности  $P(x-\varepsilon, h)$ ,  $P(x+\xi, h)$  и  $P(x, h)$  через состояния на предыдущем  $h-1$  шаге. Аналогично схеме представленной на Рис. 1 можно изобразить схемы соответствующих переходов, получить разностное уравнение для вероятности, аналогичное уравнению (2), разложить его в ряд Тейлора,

и ограничиваясь в разложении не более чем вторыми производными после ряда преобразований получим:

$$\frac{\partial P(x,t)}{\partial t} = a \frac{\partial^2 P(x,t)}{\partial x^2} - b \frac{\partial P(x,t)}{\partial x} - c \frac{\partial^2 P(x,t)}{\partial t^2}, \quad (3)$$

где:  $a = \frac{\varepsilon^2 - \varepsilon\xi + \xi^2}{\tau_0}$ ,  $b = \frac{\varepsilon - \xi}{\tau_0}$ ,  $c = \tau_0$ ,  $\frac{\partial P(x,t)}{\partial t}$  – определяет скорость изменения состояния системы с течением времени;  $\frac{\partial^2 P(x,t)}{\partial t^2}$  – описывает процесс, при котором состояния сами становятся источниками состояний (самоорганизация и ускорение);  $\frac{\partial P(x,t)}{\partial x}$  – описывает упорядоченный переход либо в состояние, когда оно увеличивается ( $\varepsilon > \xi$ ), либо, когда оно уменьшается ( $\varepsilon < \xi$ );  $\frac{\partial^2 P(x,t)}{\partial x^2}$  – описывает случайное изменение состояния (неопределенность).

## 2. Формулировка и решение краевой задачи

Считая функцию  $P(x,t)$  непрерывной, можно перейти от вероятности  $P(x,t)$  (см. уравнение (2)) к плотности вероятности  $\rho(x,t) = dP(x,t)/dx$  и сформулировать граничную задачу, решение которой и будет описывать процесс перехода между состояниями. Предположим, что необходимо контролировать в обществе долю лояльно настроенных граждан, чтобы она не опускалась ниже определенного значения (или чтобы доля отрицательно настроенных людей не превышала заданный порог для беспрепятственного распространения негативных взглядов, величина доли оппонентов должна находиться на отрезке от 0 до величины порога перехода для данной сети социальных связей (обозначим его  $m$ )). Первое граничное условие выберем исходя из следующих соображений: состояние  $x=0$  определяет полное отсутствие негативных взглядов (доля оппонентов равна 0). Сама вероятность обнаружить такое состояние может быть

отлична от 0, однако плотность вероятности, определяющую поток в состоянии  $x=0$ , необходимо положить равной 0 (состояния системы не могут выходить в область отрицательных значений (реализуется условие отражения)), т.е.:  $\rho(x, t)_{x=0} = 0$ . Второе граничное условие. Рассмотрим состояние общества с максимально возможной долей негативно настроенных индивидуумов  $x=L$  (в предельном случае  $L=1$  саму  $L$  назовем максимально возможным значением доли граждан в обществе с теми или иными настроениями,  $L$  не может быть больше 1). Сама вероятность обнаружить такое состояние будет отлична от 0. Однако плотность вероятности, определяющая поток в состоянии  $x=L$  необходимо положить равной 0 (состояния системы не могут выходить в область значений больше чем максимально возможная доля (реализуется условие отражения)), т.е.:  $\rho(x, t)_{x=L} = 0$ . Поскольку в момент времени  $t=0$  состояние системы уже может быть равно некоторому значению  $x_0$ , то начальное условие зададим в виде:

$$\rho(x, t = 0) = \delta(x - x_0) = \begin{cases} 1, & x = x_0 \\ 0, & x \neq x_0 \end{cases}$$

Так как начальное условие содержит дельта функцию, то решение для  $\rho(x, t)$  разбивается на две области при  $x > x_0$  и при  $x \leq x_0$ .

Используя методы операционного исчисления для плотности вероятности обнаружения состояния системы в одном из значений на отрезке от 0 до  $L$  можно получить следующую систему уравнений:

$$\begin{aligned} & \text{При } x \geq x_0 \\ \rho_1(x, t) &= \\ &= -\frac{2}{L} e^{-\frac{t}{2\tau_0}} e^{-\frac{(x-x_0)b}{2a}} \sum_{n=1}^{\infty} \frac{\sin\left(\pi n \frac{x_0}{L}\right) \sin\left(\pi n \frac{L-x}{L}\right)}{\cos(\pi n)} \cdot \operatorname{ch} \left( \frac{t}{\tau_0} \sqrt{\frac{\varepsilon \xi}{4a\tau_0} - \frac{\pi^2 n^2 a \tau_0}{L^2}} \right) \end{aligned}$$

$$\begin{aligned} & \text{При } x < x_0 \\ \rho_1(x, t) &= \end{aligned}$$

$$= -\frac{2}{L} e^{-\frac{t}{2\tau_0}} e^{\frac{(x-x_0)b}{2a}} \sum_{n=1}^{\infty} \frac{\sin\left(\pi n \frac{L-x_0}{L}\right) \sin\left(\pi n \frac{x}{L}\right)}{\cos(\pi n)} \cdot \operatorname{ch}\left(\frac{t}{\tau_0} \sqrt{\frac{\varepsilon\xi}{4a\tau_0} - \frac{\pi^2 n^2 a\tau_0}{L^2}}\right)$$

Если вычислить интеграл

$$P(m,t) = \int_0^{x_0} \rho_2(x,t) dx + \int_{x_0}^m \rho_1(x,t) dx,$$

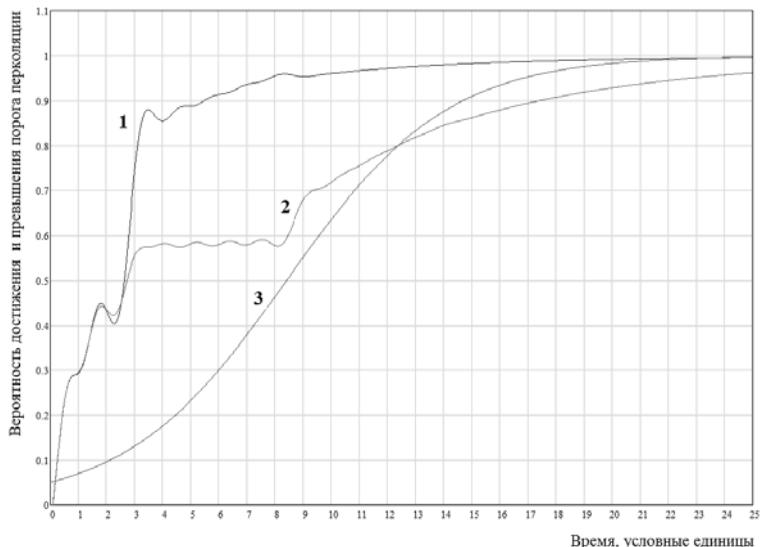
то функция  $P(m,t)$  будет задавать вероятность того, что состояние системы к моменту времени  $t$  будет находиться на отрезке от 0 до  $m$ , т.е. порог перехода  $m$  не будет достигнут. Соответственно, вероятность  $Q(m,t) = 1 - P(m,t)$  определяет то, что порог перехода  $m$  окажется к моменту времени  $t$  достигнутым.

### 3. Анализ решения краевой задачи для динамики и самоорганизации социальных состояний в обществе

Проанализируем полученные результаты. Для моделирования процесса будем считать, что доля нелояльно настроенных граждан  $x_0$  в стабильном стационарном экономическом и политическом состоянии не превышает 5% ( $x_0=0,05$ ), величину  $\tau_0$  примем равной 1 условной единице времени,  $\varepsilon=0,02$  (2%) и  $\xi=0,01$  (1%),  $L=1$  (доля всех граждан, с любыми настроениями в обществе не может быть больше 1). Результаты вычисления  $Q(m,t)$  с заданным набором параметров и различными порогами перехода для общественных сетей, представлены в графическом виде на Рис. 2. Кривая 1 на Рис. 2 построена для порога перехода равного 0,1, а кривая 2 для 0,2.

Как показывают результаты на Рис. 2, разработанная нами модель дает принципиально новые результаты, чем модель Блэкмана (кривая 3). Ход кривых 1 и 2 показывает возможность роста вероятности перехода порога перехода практически сразу после начала процесса, что, на наш взгляд, может быть связано с наличием памяти о предыдущих состояниях

системы в разработанной нами модели, и возможности описания значительной самоорганизации системы вследствие учета в дифференциальной модели второй производной по времени.



**Рис. 2.** Результаты моделирования преодоления порогов перехода негативных настроений в обществе при различных величинах начальной доли оппозиционно настроенных граждан

## Выводы

Анализ полученных данных показывает:

- чем ближе имеющаяся доля негативно настроенных граждан к порогу перехода, тем быстрее возрастает вероятность перехода;
- возможность в социальной системе нескольких скачкообразных изменений вероятности перехода через порог, когда после начала нарастания протестных настроений, через некоторое время может наблюдаться равновесие (кривая 2 на Рис. 2), а потом переход через порог;

- стохастическая динамика переходов между состояниями в сетях социальных связей демонстрирует наличие волнообразных явлений (волновую природу), что наблюдается на практике при социальных изменениях в обществе.

## Список литературы

1. *Robins G., Snijders T., Wang P., Handcock M., Pattison P.* Recent developments in exponential random graph models for social networks. // *Social Networks*, 2007. Vol. 29. P. 192–215.
2. *Airoldi E. M., Blei D. M., Fienberg S. E., and Xing E. P.* Mixed membership stochastic blockmodels // *J. Mach. Learn. Res.*, 2008. Vol. 9. P. 1981–2014.
3. *Cho Y.-S., Steeg G. V., and Galstyan A.* Co-evolution of selection and influence in social networks // In W. Burgard and D. Roth, editors, *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2011*. – San Francisco, California, USA. – AAAIPress, 2011.
4. *Du N., Wang H., and Faloutsos C.* Analysis of large multi-modal social networks: patterns and a generator // In *Proceedings of the 2010 European Conference on Machine learning and knowledge discovery in databases: Part I, ECML PKDD'10*. P. 393–408.

# Алгоритм изучения категории неопределенности в русском языке на основе корпусных данных\*

О. А. Петрушкина<sup>30</sup>, В. А. Нуриев<sup>31</sup>, Е. С. Спиридонова<sup>32</sup>

**Аннотация:** В статье рассматриваются этапы исследования показателей неопределенности в русском языке с использованием параллельного русского-французского подкорпуса, включенного в состав Национального корпуса русского языка (НКРЯ). Приводится методология, используемая для их изучения, а также предлагается оригинальная структура описания данных языковых единиц.

**Ключевые слова:** категория неопределенности; параллельный корпус; корпусная лингвистика

## Введение

Данная статья посвящена описанию алгоритма, разработанного для изучения категории неопределенности в русском языке. В каждом из разделов статьи описывается его отдельный этап.

В этой связи возникает вопрос, необходимо ли сейчас изучать категорию неопределенности в русском языке и какие методы исследования при этом использовать. Категория неопределенности в русском языке отличается от других языков: неопределенность является одной из «семантических доминант» русского языка и непосредственно связана с языковой картиной мира, которую она формирует.

Русские показатели неопределенности крайне разнообразны

---

\*Работа выполнена при финансовой поддержке РФФИ (проект № 14-06-31358\_мол\_а).

<sup>30</sup> ФИЦ ИУ РАН (ИПИ РАН); 1987 г.р.; olga.petrouchkina@gmail.com

<sup>31</sup> ИЯз РАН; ФИЦ ИУ РАН (ИПИ РАН); 1980 г.р.; к.филол.н.; nurieff.v@gmail.com

<sup>32</sup> ФИЦ ИУ РАН (ИПИ РАН); 1991 г.р.; spiridliza@rambler.ru

разны и могут выражать тончайшие нюансы смысла. Соответственно, они не отличаются однородностью грамматических признаков. Это значительно затрудняет классификацию, а также их формализованное выявление. Следует отметить, что к проблеме изучения маркеров неопределенности в русском языке уже неоднократно обращались. Однако эти исследования проводились в докорпусную эпоху, и потому их результаты требуют верификации. Другими словами, уже существующие типологии нуждаются в дополнении, что требует более детального изучения русских маркеров неопределенности.

Говоря о методах исследования, необходимо подчеркнуть, что корпусная методика сбора материала для изучения категории неопределенности используется впервые. Она позволяет обрабатывать значительные объемы информации, которые дают возможность подтвердить и дополнить полученные ранее результаты. В данной работе русские показатели неопределенности исследуются на фоне французского языка, так как французские соответствия позволяют показать выраженность степени неопределенности той или иной русской единицы.

## **1. Этапы исследования**

### **1.1. Пополнение русско-французского подкорпуса НКРЯ**

Пополнение корпуса является одним из важных этапов исследования: 1) в силу необходимости расширения корпусного материала; 2) из-за того, что до недавнего момента в русско-французском подкорпусе были представлены только классические произведения при полном отсутствии современных. В рамках нашего проекта подкорпус был пополнен текстом документально-художественного произведения С. Алексиевич «Время секонд хэнд» (М.: «Время», 2013) и его перевода на французский язык (Alexievitch S. «La fin de l'homme rouge ou le temps du désenchantement», Actes Sud, 2013). Подкорпус был пополнен на 342774 словоупотребления.

Для этого использовалась методика, разработанная в ходе проекта «Контрастивные корпусные исследования русских

и французских глагольных категорий в поливариантных параллельных текстах» [1]. Она включает в себя шесть этапов: отбор параллельных текстов; оцифровка параллельных текстов; выравнивание текстов; метатекстовая разметка текстов; морфологическая разметка текстов и загрузка их в корпус. Выравнивание и разметка текста осуществлялись при помощи программы EuclidParalleltextaligner. Сейчас текст находится в свободном доступе для работы на сайте НКРЯ. Это позволяет производить поиск показателей неопределенности и анализировать особенности их функционирования не только в языке классических произведений, но и в современных текстах.

## **1.2. Совершенствование методики поиска и отбора показателей неопределенности**

В качестве основы для типологии показателей неопределенности была использована классификация, предложенная Е. В. Падучевой в книге 1985 г. «Высказывание и его соотносительность с действительностью (Референциальные аспекты семантики местоимений)» [3] и в статье 1996 г. «Неопределенность как семантическая доминанта русской языковой картины мира».

Однако при анализе примеров, полученных при помощи корпусных данных, возникла трудность в определении необходимого минимального контекста для изучения рассматриваемых языковых единиц. Так, требуется разная протяженность контекста при изучении семантики собственно неопределенного местоимения кто-то, показателя эпистемической модальности должно быть и неопределенно-личных предложений. В результате корпусно-ориентированного сопоставительного анализа было выявлено, что минимальным анализируемым контекстом могут быть два выровненных параллельных предложения, где есть конкретный показатель неопределенности и его эквивалент (в нашем случае – французский).

### 1.3. Инвентаризация показателей неопределенности

Инвентаризация исследуемых единиц проходила следующим образом. Методом сплошной выборки по параллельному корпусу было отобрано 1500 контекстов, в которых имелись единицы русского языка, выражающие значение неопределенности. Затем, на основе указанной выше классификации, была построена типология этих маркеров:

#### 1) Неопределенные местоимения:

собственно неопределенные местоимения (*кто-то, какой-то, чей-то и т.д.*); слабоопределенные местоимения (*кое-какой, кое-кто, некто, некий, некоторые и т.д., один*); нереферентные местоимения (*какой-нибудь, кто-нибудь, кто-либо, иной*).

#### 2) Неопределенные наречия:

собственно неопределенные наречия (*когда-то, некогда, где-то и т.д.*); слабоопределенные наречия (*однажды, как-то раз, кое-где, кое-когда и т.д.*); нереферентные наречия (*когда-нибудь, когда-либо, где-нибудь, куда-либо и т.д.*).

3) Показатели эпистемической модальности (*должен, должно быть, вроде (бы), не то... не то, то ли... то ли, или; чуть ли не; как будто, как бы; что ли; что-то; почему-то и др.*).

#### 4) Неопределенно-личные предложения:

с неопределенным субъектом, представляющим открытое множество (*Я совсем не хочу, чтобы мою книжку читали просто ради забавы*); с неопределенным субъектом, представляющим закрытое множество (*Учительница сказала нам, что нас будут фотографировать*); с неопределенным единичным субъектом (*Тебя прощают. Не будут больше подавать платок*).

### 1.4. Разработка универсальной структуры описания

Для параметризованного анализа и описания каждого конкретного показателя неопределенности было необ-

ходимо разработать универсальную структуру, которая может быть применена к любому типу показателей неопределенности. Также было необходимо уточнить терминологический аппарат: в частности, ввести такие параметры как, например, точка фокализации (отображение точки зрения коммуникантов в высказывании [2]), степень неопределенности и т.д. Далее рассмотрим описание русских маркеров неопределенности на примере собственно неопределенного местоимения *кто-то*.

**Класс:** неопределенные местоимения. **Подкласс:** собственно неопределенные местоимения. **Семантическая функция:** референт неизвестен ни говорящему, ни адресату.

**Фокализация:**

- внутренняя (сфера действия неопределенности ограничивается рамками представленной в предложении/контактных предложениях коммуникативной ситуации).

Пример 1:

*Звонок. Кто-то только что явился в магазин.* [М. Булгаков Белая гвардия (1924)]

*La sonnette. Quelqu'un venait d'entrer dans la boutique.* [Mikhail Boulgakov. La garde Blanche (Claude Ligny, 1970)]

- нулевая (сфера действия неопределенности выходит за рамки представленной в предложении/контактных предложениях коммуникативной ситуации).

Пример 2:

*Река тоже присмирела; немного погода и в ней вдруг кто-то плеснул еще в последний раз, и она стала неподвижна.* [И.А. Гончаров. Обломов (1848-1859)]

*Et celle-ci s'assagit à son tour: un dernier clapotement, et elle est immobile.* [Ivan Goncharov, Oblomov, (Arthur Adamov 1959)]

*... et de la rivière: endormie: peu après on y entend un dernier clapotis, puis elle s'immobilise.* [Ivan Goncharov, Oblomov, (Luba Jurgenson 1988)]

В примере 2 показано, что русский показатель неопределенности *кто-то* удаляется из перевода, не получая реализации в качестве полнозначной лексемы, а контакт-

ный русский глагол трансформируется во французское существительное (*плескать* – *clapotement, clapotis*), при котором ставится неопределенный артикль.

Реконструируя ситуацию из данного примера, можно заметить, что *кто-то* вводит неопределенность персонажа, но персонажа второстепенного, не несущего дальнейшей функциональной нагрузки в сюжетной развертке. Соответственно признак неопределенности не требует обязательной акцентуации и интенсификации, потому маркируется грамматическими средствами. По той же причине при переводе *кто-то плеснул* в обоих случаях нивелируется признак одушевленности, характерный для русского *кто-то*. Таким образом, введение местоимения *кто-то* задает нестабильность точки фокализации (в отдельных случаях, как в последнем примере, ее идентификация вообще затруднена).

Семантические особенности: местоимение *кто-то*, как правило, относится к настоящему и прошедшему времени, может выражать (в зависимости от контекста) идею единичности и множественности, на периферии значение может сближаться с *кто-нибудь*, может выполнять интродуктивную функцию. Синтаксические особенности: местоимение *кто-то* замещает функцию субъекта, может употребляться как в главном, так и в придаточном предложении. Степень неопределенности: высокая.

Лингвоспецифичность собственно неопределенного местоимения *кто-то* заключается в том, что оно может соотноситься как с закрытым/открытым множеством, так и с единичным референтом, что исключается, например, в отношении его французского прямого переводного эквивалента *quelqu'un*.

## Выводы

Данное исследование направлено на внутреннюю классификацию показателей неопределенности в русском языке, а также на установление универсального инвентаря языковых

значений, что является одной из самых актуальных задач современной типологии (перспективность исследований такого рода отмечена, например, в работах [4; 5]).

В рамках данного исследования продолжается разработка идеи о существовании семантических доминант языка. Таким образом, оно является сравнительным исследованием в области культурно-специфичной языковой семантики (Исследованиями в области культурно-специфичной языковой семантики активно занимается известный лингвист А. Вежбицкая (см., например, [6])).

Корпусное исследование маркеров неопределенности в русском языке на фоне французского проводится впервые и дает возможность параллельного анализа данных показателей в двух языках. В дальнейшем это поможет уточнить семантический объем маркеров неопределенности в разных языках и установить семантические зоны, в которых их значения пересекаются.

## Список литературы

1. Бунтман Н. В., Зализняк А. А., Зацман И. М., Кружков М. Г., Лоцилова Е. Ю., Сичинава Д. В. Информационные технологии корпусных исследований: принципы построения кросслингвистических баз данных // Информатика и ее применения, 2014. Т.8, №.2. С. 98–110.
2. Женетт Ж. Фигуры. – М.: Изд-во им. Сабашниковых, 1998. Т.2. – 472 с.
3. Падучева Е. В. Высказывание и его соотносительность с действительностью (Референциальные аспекты семантики местоимений). – М.: Наука, 1985. – 272 с.
4. Плузьян В. А. Общая морфология. Введение в проблематику. – М.: Эдиториал УРСС. 2000. – 384 с.
5. Плузьян В. А. Введение в грамматическую семантику: грамматические значения и грамматические системы языков мира. – М.: Издательство РГГУ, 2011. – 672 с.
6. Goddard Cl., Wierzbicka A. Words and Meanings. Lexical Semantics Across Domains, Languages, and Cultures. – Oxford: Oxford University Press, 2014. – 336 p.

# Методика построения моноэквиваленций в надкорпусной базе данных коннекторов

Н.А. Попкова<sup>33</sup>, О.Ю. Инькова<sup>34</sup>, И.М. Зацман<sup>35</sup>,  
М.Г. Кружков<sup>36</sup>

**Аннотация:** Рассматривается методика формирования надкорпусной базы данных (НБД) коннекторов как лингвистического информационного ресурса, создаваемого в качестве надстройки к электронным подкорпусам параллельных текстов Национального корпуса русского языка. Наполнение НБД основано на принципе формирования моноэквиваленций, каждая из которых представляет собой упорядоченную пару, состоящую из изучаемой языковой единицы и ее перевода, а также из их формализованного описания, включающего параметры, характеризующие ее функционирование. Коннекторы, т.е. языковые единицы со связующей функцией, могут быть по своему составу одно- и многокомпонентными. Описание как последних в целом, так и их составляющих основано на использовании кортежей нескольких категорий, рассматриваемых в статье. Показано, что НБД коннекторов и предлагаемый подход к категоризации их моноэквиваленций существенно расширяют спектр аспектов изучения этой языковой единицы, обеспечивают накопление новых и развитие существующих знаний о коннекторах.

**Ключевые слова:** компьютерная лингвистика; информационные ресурсы; контрастивная лингвистика; коннекторы; надкорпусная база данных; русский язык; французский язык; параллельные тексты

## Введение

Вопросы, обсуждаемые в статье, связаны с разработкой методологии и технологий создания информационных ресурсов

---

<sup>33</sup> ФИЦ ИУ РАН (ИПИ РАН); 1992 г.р.; natasha\_\_popkova@mail.ru

<sup>34</sup> ФИЦ ИУ РАН (ИПИ РАН); 1965 г.р.; д.филол.н.; o.inkova@bluewin.ch

<sup>35</sup> ФИЦ ИУ РАН (ИПИ РАН); 1952 г.р.; д.т.н.; izatsman@ya.ru

<sup>36</sup> ФИЦ ИУ РАН (ИПИ РАН); 1975 г.р.; magnit75@yandex.ru

в виде электронных корпусов и баз данных. Их использование в области контрастивных лингвистических исследований позволяет получать новые фундаментальные знания в компьютерной лингвистике, теории перевода, корпусной лингвистике, теории целенаправленной генерации знаний и когнитивной информатике.

Статья посвящена описанию методики создания нового вида лингвистических информационных ресурсов, который получил название «надкорпусные базы данных» (НБД) [1]. Слово «надкорпусная» используется потому, что НБД создается как надстройка к электронным корпусам текстов. При анализе текстов, хранящихся в корпусах любого типа, исследователю необходимо сохранять результаты работы не столько в бумажной, сколько в электронной форме, и «в такой ситуации возникает необходимость в разработке специализированной технологии, позволяющей последовательно обрабатывать все полные тексты корпусов (или подкорпусов) и фиксировать в отдельной БД каждое появление исследуемого языкового явления» [1], а также хранить и искать его формализованное описание по широкому спектру поисковых лексических и грамматических признаков как в одном, так и в нескольких языках.

На сегодняшний день в ФИЦ ИУ РАН разработано три НБД: для личных глагольных категорий, лингвоспецифичных единиц и коннекторов русского языка. Их описанию посвящены работы [2, 3, 4]. В двух последних подробно рассмотрена методика формирования БД глагольных категорий.

Целью данной работы является описание методики формирования НБД коннекторов, содержащей формализованные описания исследуемых языковых единиц русского языка и их переводных функционально эквивалентных фрагментов (ФЭФ, введен Д.О. Добровольским) во французском языке [5].

## **1. Принцип построения НБД коннекторов**

НБД коннекторов предназначена для их формализованного описания, в том числе как многокомпонентных языковых

единиц. Коннектор – это «языковая единица, функция которой состоит в выражении типа отношений – логико-семантических, иллюкутивных<sup>37</sup>, структурных, – существующих между двумя соединенными с ее помощью компонентами, имеющими предикативный характер, выраженными имплицитно или эксплицитно» [6]. Их лексико-грамматическая природа может быть различна. Коннектор, кроме того, может состоять из одного (*хотя, притом, кстати*) и большего числа слов (*не только..., но и; скорее..., чем...*); первые являются по своему составу однокомпонентными (пример 1), вторые – многокомпонентными (пример 2).

Пример 1. [Гончаров. Обломов]:

*Нет, у него чернильница полна чернил, на столе лежат письма, бумага, даже гербовая, **притом** исписанная его рукой.*

Пример 2. [Гончаров. Обломов]:

*Захар не старался изменить **не только** данного ему богом образа, **но и** своего костюма, в котором ходил в деревне.*

С лингвистической точки зрения интерес представляют как многокомпонентные коннекторы в целом, так и их составные части, причем в некоторых случаях части коннектора могут в другом контексте сами по себе являться коннекторами, например, *только* (ср. его употребление в примере 2). Поэтому основополагающим принципом наполнения НБД коннекторов является следующий: многокомпонентный коннектор как языковая единица описывается сначала в целом, а затем описываются по отдельности составляющие его элементы.

Информация о функционировании коннекторов фиксируется в НБД в виде моноэквиваленций (МЭ) или кортежей, то есть упорядоченных пар, состоящих из исследуемой языковой единицы и ее ФЭФ, которые сопровождаются их формализованными описаниями.

---

<sup>37</sup> Т.е. устанавливаемых на уровне речевых актов. Так, в высказывании *Он не пришел, потому что заболел* коннектор *потому что* устанавливает логико-семантическое отношение причины на уровне двух ситуаций, описанных в главном и в придаточном предложении. А в высказывании *Позвони ему, пожалуйста, потому что он может обидеться* тот же коннектор устанавливает то же отношение, но на уровне речевого акта, совершенного в главном предложении (я тебя **прошу** позвонить), и вводит причину этой просьбы (иначе он обидится).

В НБД уже построена МЭ для коннектора из примера 2 и следующего его перевода на французский язык:

Пример 3: *Zakhar n'avait rien fait pour changer l'apparence que Dieu lui avait donnée ni le costume qu'il avait porté à la campagne.*

Для коннектора *не только..., но и* и этого его перевода в рассматриваемой НБД построен кортеж следующего вида: (**не только..., но и; ni**). Эта пара лексических единиц русского и французского языка была сформирована в процессе анализа предложений, которые приведены в примерах 2 и 3. В НБД этот кортеж сопровождается набором полей с дополнительной информацией, которые будут описаны в разделе 2.

## 2. Методика создания МЭ

Первый этап построения МЭ – это поиск в НБД по шаблонам, которые формируются пользователем для каждого исследуемого коннектора. В результате поиска по сформированному запросу из заданного массива параллельных текстов отбираются все пары предложений, содержащие исследуемый коннектор. После этого лингвист анализирует каждую найденную пару предложений в отдельности. В процессе поиска отбираются также и те предложения, которые содержат полифункциональные лингвистические единицы, т.е. те, которые могут выступать и не в функции коннектора. Например, лексическая единица *и* может быть коннектором, если она выступает в роли союза, а может им не быть, если она функционирует как частица.

На втором этапе построения МЭ для коннекторов в поля НБД вводится дополнительная информация, характеризующая данное употребление описываемой языковой единицы. Отмечаются все составляющие коннектора русского языка, определяется вид коннектора согласно его принадлежности к тому или иному структурному кластеру (группе коннекторов). С целью группировки коннекторов был разработан экспериментальный список кластеров, которые объединяют коннекторы

как с общей составляющей, так и с некоторым общим признаком.

Например, кластер «При» содержит следующие коннекторы с их общей составляющей «при»: *при <расстояние> том, при <расстояние> чем, при <расстояние> этом, притом, причем*. Кластер «Сочетания» был создан для учета наиболее употребляемых сочетаний коннекторов: *а <расстояние> иначе, а <расстояние> впрочем* и т.д., т. е. по признаку частотности использования одновременно нескольких коннекторов. Эта типология находится на начальной стадии своего формирования. В процессе ее развития число оснований классификации коннекторов может быть увеличено.

На третьем этапе построения МЭ определяется ФЭФ во французской фразе, который соответствует исследуемому коннектору русского языка. Французский ФЭФ коннектора определяется в результате семантического анализа русского предложения и сопоставления с ним предложения французского перевода. Не всегда удается вычленить абсолютно точно эквивалентную ФЭФ, что отмечается в полях формализованного описания коннекторов с помощью признаков МЭ. Опыт наполнения НБД коннекторов показал, что множество ФЭФ французского языка, соответствующих одному коннектору русского языка, может пополняться за счет вариантов перевода, отражающих предпочтения переводчиков. Например, на данный момент для коннектора *причем* зафиксировано 8 ФЭФ в форме коннекторов (*et, de plus, en même temps, d'ailleurs, pendant que, alors, puis, à cette occasion*), которые существенно разнятся по частотности их встречаемости в корпусе.

Сформированное в процессе анализа коннекторов множество их ФЭФ также кластеризовано с учетом того, что не всегда коннектор переводится коннектором. Было сформировано три кластера (термины заимствованы из [7]): конвергентный перевод (для случаев, когда коннектор русского языка переведен коннектором французского языка), дивергентный перевод (для случаев, когда коннектор русского языка переведен другой

языковой единицей или конструкцией, или не переведен вообще) и кластер с ФЭФ, чей статус лингвистам еще предстоит определить.

Далее для сформированной пары (*коннектор: ФЭФ*), например, (**не только...**, **но и: ni**) заполняются поля дополнительных признаков. Например, в МЭ, построенной на базе примеров 2 и 3, для коннектора русского языка были заполнены следующие поля. Поле «Статус» для примера 2 содержит помету «CNT», которая означает, что анализируемая лексическая единица русского языка в данном контексте является коннектором. Поле «Порядок» содержит помету «CNT<sub>1</sub> p CNT<sub>2</sub> q», что указывает на взаимное расположение частей коннектора и частей предложения, которые они связывают (для примера 2 составные элементы пометы CNT<sub>1</sub> p CNT<sub>2</sub> q обозначают: CNT<sub>1</sub> – **не только...**, p – *данного ему богом образа*, CNT<sub>2</sub> – **но и**, q – *своего костюма, в котором ходил в деревне*).

Формализованное описание ФЭФ для примера 3 включает следующие поля. Поле «Отношения» содержит помету «соединительные», т.е. этот коннектор французского языка выражает соединительные отношения между частями предложения. Поле «Позиция» содержит помету «начальная», т.е. коннектор занимает начальную позицию во фрагменте французского текста, который он маркирует. Поле «Порядок» содержит помету «p CNT q», где p – *l'apparence que Dieu lui avait donnée*, CNT – **ni**, q – *le costume qu'il avait porté à la campagne*. Поле «Статус» содержит помету «CNT», это означает, что лексическая единица французского языка является коннектором.

Таким образом, в формализованном описании МЭ (*не только...*, *но и: ni*), построенной для примеров 2 и 3, отражены особенности структуры коннектора русского языка и его ФЭФ, а также особенности их функционирования в тексте.

Формализованное описание МЭ содержит ряд параметров, характеризующих сформированный кортеж в целом: необходимость дополнительной экспертной оценки, расширения контекста для более подробного семантического анализа и описания коннектора, а также поле «nota bene» (для пометки особо

интересных, по мнению исследователя, случаев). Лингвист принимает решение о необходимости проставления этих помет в каждом конкретном случае. В поле «Комментарии» и лингвист, формирующий МЭ, и эксперт могут оставлять свои замечания в текстовой форме.

### 3. Категории кортежей

Как было отмечено выше, кортеж, построенный для коннектора, представляет собой упорядоченную пару, содержащую эту русскую языковую единицу и соответствующий ей фрагмент французского текста. Рассмотренный пример кортежа (**не только...**, **но и; ni**) относится только к одной из 9 категорий, определенных в концепции формирования НБД. В процессе сопоставительного анализа параллельных русских и французских текстов, полученных из Национального корпуса русского языка (НКРЯ), в настоящее время в НБД коннекторов лингвистами формируются русско-французские соответствия в виде трех из девяти категорий двухместных кортежей:

1. (с, Fс), где с – коннектор русского языка, Fс – его ФЭФ на французском языке;
2. (b, Fb), где b – блок коннектора (его неэлементарная составная часть), Fb – его ФЭФ (число блоков и необходимость их описания в НБД определяется лингвистом);
3. (e, Fe), где e – элемент коннектора (его элементарная составная часть), Fe – его ФЭФ (число описываемых элементов определяется лингвистом).

В процессе описания каждого многокомпонентного коннектора могут использоваться кортежи всех трех перечисленных категорий, но необходимость описания компонента коннектора кортежа определяется лингвистом. Обязательным является построение только кортежа первой категории, т.е. для всего многокомпонентного коннектора.

Приведем два примера коннекторов и их переводов, для описания структуры которых может понадобиться использование кортежей нескольких категорий.

Пример 4. [Достоевский. Преступление и наказание]:

...**но все же** несколько оскорбленного в своем достоинстве человека,  
**и притом** твердо решившегося потребовать объяснений.  
...*mais, néanmoins, un peu offensé dans sa dignité et, en même temps,*  
*fermement décidé à exiger des explications.*

В примере 4 для обоих коннекторов *но все же* и *притом* лингвист может принять решение построить кортежи первой и третьей категории. Хотя блок *все же* коннектора *но все же* не является элементарным по своему составу, т.е. состоит, в отличие от *притом*, из двух компонентов, не имеет смысла разбивать *все же* на элементарные единицы, поскольку элемент *все* выполняет связующую функцию только в сочетании с другими элементами.

Пример 5. [Гончаров. Обломов]:

...мягкость, которая была господствующим и основным выражением, **не** лица **только**, **а** всей души;  
...*L'expression de douceur, qui dominait non seulement le visage mais aussi l'âme.*

Для многокомпонентного коннектора из примера 5 *не... только, а* могут быть построены кортежи всех трех категорий, поскольку он состоит из разложимого блока *не только*, для которого может быть построен кортеж второй категории, и из элементарной единицы *а*, для которой может быть построен кортеж третьей категории (также при необходимости могут быть построены кортежи для элементов, составляющих блок *не только*).

После построения лингвистом вышеописанных двухместных кортежей, автоматически, т.е. программным способом без участия лингвистов, могут быть сформированы двухместные русско-французские кортежи еще трех категорий, в которых второй компонент является множеством ФЭФ из разных переводов одного и того же оригинального текста:

1.  $(с, \{F_{c1}, \dots, F_{cn}\})$ , где  $с$  – коннектор русского языка,  $F_{c1}$  – его ФЭФ в первом переводе,  $\dots$ ,  $F_{cn}$  – его ФЭФ в  $n$ -ом переводе;

2.  $(b, \{Fb_1, \dots, Fb_n\})$ , где  $b$  – блок коннектора,  $Fb_1$  – его ФЭФ в первом переводе,  $\dots$ ,  $Fb_n$  – его ФЭФ в  $n$ -ом переводе;
3.  $(e, \{Fe_1, \dots, Fe_n\})$ , где  $e$  – неделимый элемент коннектора,  $Fe_1$  – его ФЭФ в первом переводе,  $\dots$ ,  $Fe_n$  – его ФЭФ в  $n$ -ом переводе.

В перспективе также планируется провести сопоставительный анализ параллельных французских и русских текстов, состоящих из выровненных по предложениям оригинальных произведений на французском языке и их переводов на русский, полученных из НКРЯ. Это даст возможность лингвистам сформировать в НБД коннекторов следующие францужско-русские соответствия еще трех категорий в виде двухместных кортежей:

1.  $(Sc, c)$ , где  $c$  – многокомпонентный коннектор русского языка,  $Sc$  – «стимул» французского языка, обусловивший появление этого коннектора в переводе на русский язык;
2.  $(Sb, b)$ , где  $b$  – блок коннектора русского языка,  $Sb$  – «стимул» французского языка, обусловивший появление этого блока коннектора в переводе на русский язык;
3.  $(Se, e)$ , где  $e$  – неделимый элемент коннектора русского языка,  $Se$  – «стимул» французского языка, обусловивший появление этого элемента коннектора в переводе на русский язык.

Одновременное формализованное описание в НБД не только всего многокомпонентного коннектора как единой языковой единицы, но и его составляющих позволяет, с одной стороны, фиксировать все возможные блоки и элементы, составляющие коннекторы, характеризующиеся высокой степенью вариативности, а с другой стороны, даст возможность отслеживать, в какие многокомпонентные коннекторы может входить тот или иной блок или его элемент.

## Выводы

Работа по наполнению НБД коннекторов согласно описанной методике продемонстрировала, что функциональность НБД существенно дополняет функциональность электронных корпусов параллельных текстов и значительно расширяет возможности их использования в процессе контрастивного лингвистического анализа.

Первым таким преимуществом можно считать тот факт, что, в отличие от корпусов параллельных текстов, в которых возможна лишь работа по поиску и отбору предложений, содержащих изучаемую языковую единицу, НБД позволяет исследователю также самому описывать эту единицу в виде набора полей, исходя из своих лингвистических задач (в данном случае описание коннектора как многокомпонентной единицы), создавая формализованное описание семантических и синтаксических свойств этой единицы, потом сохранять его, что позволяет затем многократно производить поиск, используя созданное описание. Второе преимущество заключается в возможности сохранять в НБД формализованное описание не только коннектора русского языка, но и описание его ФЭФ во французском языке, которое также имеет научную ценность, так как в электронной форме сохраняются результаты контрастивного анализа, и становится возможным двуязычный лексико-грамматический поиск исследуемых языковых единиц.

Наконец, статистические данные НБД о частности использования коннекторов и их характеристиках являются верифицируемыми. Использование функции НБД под названием «Просмотр статистики» дает сведения о числе моноэквиваленций, содержащих определенное сочетание коннектор–ФЭФ, и это число является гиперссылкой, по которой становятся наглядно видны соответствующие МЭ, их состав, а также набор полей, содержащих формализованное описание семантических и синтаксических свойств коннекторов.

Все перечисленные функции НБД являются достаточным основанием для того, чтобы выделить их в отдельную ка-

тегорию лингвистических информационных ресурсов, обладающих принципиально новыми функциями и возможностями по сравнению с электронными корпусами параллельных текстов.

## Список литературы

1. *Кружков М. Г.*, Информационные ресурсы контрастивных лингвистических исследований: электронные корпуса текстов // Системы и средства информатики, 2015. Т. 25. Вып. 2. С.140–159.
2. *Зализняк А. А., Зацман И.М., Инькова О.Ю., Кружков М.Г.* Надкорпусные базы данных как лингвистический ресурс // Корпусная лингвистика - 2015: Труды 7-й Международной конференции – СПб.: СПбГУ, 2015. С.211–218.
3. *Бунтман Н. В., Зализняк А. А., Зацман И. М., Кружков М. Г., Лощилова Е. Ю., Д. В. Сичинава Д. В.* Информационные технологии корпусных исследований: принципы построения кросс-лингвистических баз данных // Информатика и ее применения, 2014. Том 8. Выпуск 2. С. 98–110.
4. *Kruzhkov M.G., Buntman N.V., Loshchilova E.Ju., Sitchinava D.V., Zalizniak A. A., Zatsman I.M.* A database of Russian verbal forms and their French translation equivalents, Компьютерная лингвистика и интеллектуальные технологии (По материалам ежегодной Международной конференции «Диалог-2014»). Том 13(20). 2014. С. 284–297.
5. *Добровольский Д.О., Кретов А.А., Шаров С.А.* Корпус параллельных текстов // Научная и техническая информация. Сер.2 Информационные процессы и системы, 2005. № 6. С.16–27.
6. *Инькова-Манзотти О. Ю.* Коннекторы противопоставления во французском и русском языках: сопоставительное исследование, М: Информэлектро, 2001.
7. *Johansson S.* Seeing through Multilingual Corpora. On the use of corpora in contrastive studies. – Amsterdam: John Benjamins, 2007.

# Обнаружение границ препятствий в видеопотоке, полученном с фронтальной стереопары

Н.В. Решетняк<sup>38</sup>, В.В. Постников<sup>39</sup>

**Аннотация:** В работе рассматривается задача обнаружения объектов, препятствующих движению транспортного средства. Изложен быстрый метод обнаружения препятствий на основе выделения вертикальных границ объектов. Предложена и экспериментально проверена функция сравнения граничных пикселей изображения.

**Ключевые слова:** стереозрение; обнаружение препятствий; функция сравнения.

## Введение

В последние годы наблюдается возросший интерес к созданию систем управления транспортным средством с помощью компьютерного зрения [1, 2]. Одной из ключевых задач беспилотного управления является обнаружение препятствий дорожной сцены. К таким препятствиям могут относиться как статические объекты (дорожные знаки, разделительные полосы, столбы со светофорами), так и динамические (пешеходы, автомобили). Обнаружение препятствий и своевременное реагирование (остановка или объезд) определяют экономическую целесообразность и общественную безопасность использования беспилотного транспортного средства [1].

## 1. Обзор существующих методов

Существует два подхода к анализу дорожной сцены [3].

---

<sup>38</sup> МФТИ; 1992 г.р.; nikita.reshetnyak@phystech.edu

<sup>39</sup> ФИЦ ИУ РАН (ИСА РАН); к.н.; vassili.postnikov@gmail.com

Во-первых, полный анализ видимой дорожной сцены. Например, методы на основе построения полной карты глубин [4]. Такой способ анализа дорожной сцены требует больших вычислительных ресурсов, что в свою очередь затрудняет использование метода в режиме реального времени.

Во-вторых, анализ локальных признаков сцены. Данный подход позволяет построить только разреженную карту глубин, но методы на основе локальных признаков сцены более устойчивы к помехам и могут быть без проблем применены для определения препятствий в режиме реального времени.

В работе [3], в качестве локальных признаков предлагается использовать вертикальные границы и углы. На основе этих признаков авторами был разработан способ деления границ на два класса.

- Границы объектов, лежащих в плоскости земли. К таким объектам относятся границы дороги, дорожная разметка и дефекты дорожного полотна.

- Границы объектов препятствующих движению автотранспорта.

Алгоритм работает на ректифицированных изображениях [5] стереопары. Пусть объект лежит в плоскости земли. Выделим граничные пиксели этого объекта. Тогда для любого граничного пикселя на левом изображении можно получить положение этого пикселя на правом изображении по формуле (1) [3].

$$d(x, y) = \frac{B}{H} f_x \left( \frac{y}{f_y} \cos \alpha + \sin \alpha \right), \quad y > f_y \tan \alpha, \quad (1)$$

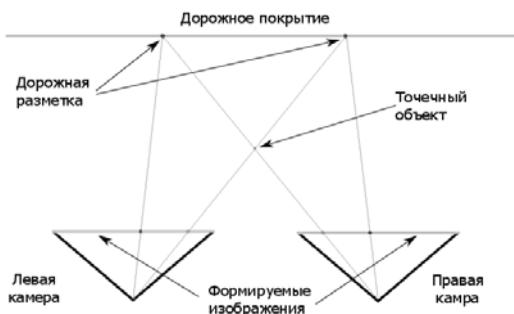
где  $d(x, y)$  – смещение (disparity) граничного пикселя объекта, лежащего в плоскости земли;  $(x, y)$  – координаты пикселя на левом изображении;  $B$  – база стереопары (расстояние между двумя камерами);  $H$  – высота крепления камеры;  $\alpha$  – угол наклона камеры к горизонту;  $f_x$  и  $f_y$  – фокусные расстояния камеры.

Таким образом, алгоритм строит бинарную маску объектов, лежащих в плоскости земли, как пересечение границ на левом и на правом кадре с учетом смещения каждого пикселя.

Все остальные границы относятся к границам препятствий.

## 2. Функция сравнения

Описанный алгоритм [3] имеет один недостаток. Граничный пиксель препятствия будет ошибочно отнесен к дорожному покрытию, если он совпадет с граничным пикселем другого объекта. На Рис. 1 показана плоскость, образованная главными осями камеры. Для простоты будем считать препятствие и дорожную разметку точечными объектами. Из рисунка видно, что точечное препятствие заслоняет правую точку разметки на левой камере и левую точку разметки на правой камере, в результате применения алгоритма объект будет утерян, так как границы объектов будут в точности совпадать с границами, если бы объект отсутствовал.



**Рис.1.** Пример препятствия, которое будет отнесено к дорожному полотну при сравнении границ.

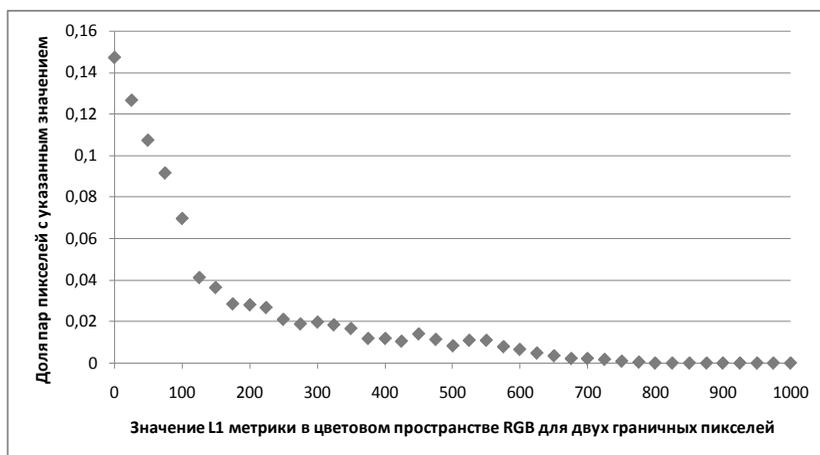
В качестве числовой функции сравнения граничных точек предлагается использовать L1 метрику (2) в цветовом пространстве RGB. То есть, вместо бинарного сравнения границ будем сравнивать исходные значения пикселей.

$$\|X - Y\| = |R_X - R_Y| + |G_X - G_Y| + |B_X - B_Y|, \quad (2)$$

где  $X$  – граничный пиксель на левом изображении;  $Y$  – образ пикселя  $X$  на правом изображении (если  $X$  лежит в плоскости земли);  $(R_i, G_i, B_i)$  – значение пикселя  $i$  в цветовом пространстве RGB.

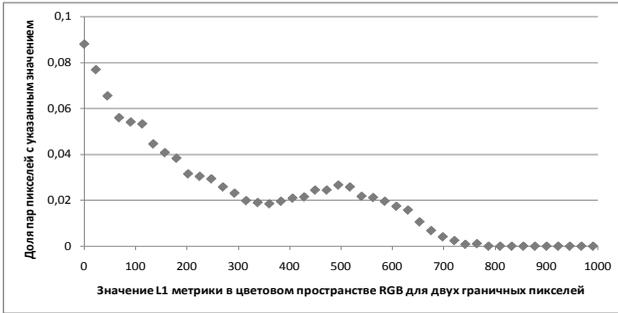
### 3. Результаты экспериментов

Мы исследовали распределение значений предложенной метрики. На Рис. 2 показано распределение значений для сцены без препятствий, полученные значения монотонно убывают.

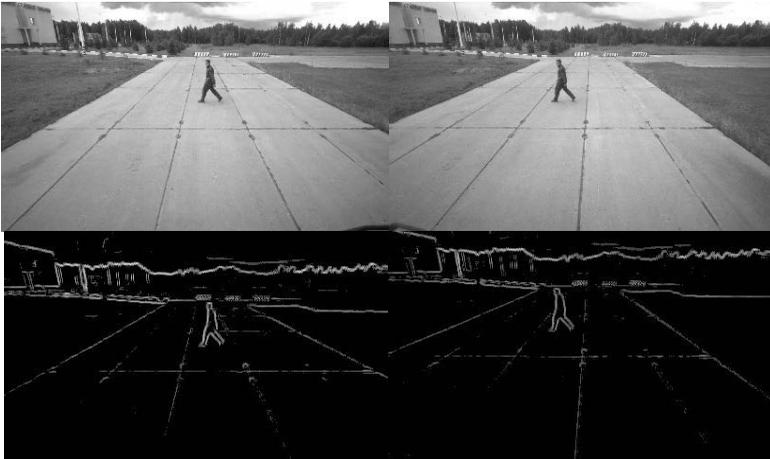


**Рис. 2.** Распределение метрики (2) при отсутствии препятствий.

Распределение значений (2) при наличии препятствий представлено на Рис. 3. Из распределения следует, что существует порог, разделяющий границы препятствия и границы, лежащие в плоскости земли. Экспериментально был установлен порог отсечения равный 350.



**Рис.3.** Распределение метрики (2) при наличии препятствий.



**Рис.4.** Пример изображений стереопары и результат выделения границ



**Рис. 5.** Выделенные границы препятствий, расположенных выше плоскости дороги

На Рис. 4 приведен пример исходных изображений стереопары и выделенные вертикальные границы. Результат работы алгоритма с предложенной функцией расстояния показан на Рис. 5.

## **Выводы**

В работе был рассмотрен алгоритм выделения границ препятствий по двум изображениям стереопары. Была предложена метрика сравнения граничных пикселей изображения, устойчивая к заслонению и неверной классификации границ. Было проанализировано распределение значений функции сравнения граничных пикселей и экспериментально показана ее эффективность.

## **Список литературы**

1. *Distner M. et al.* City safety—a system addressing rear-end collisions at low speeds //Proc. 21st International Technical Conference on the Enhanced Safety of Vehicles. – 2009. – №. 09-0371.
2. *Gavrila D. M. et al.* Real time vision for intelligent vehicles //Instrumentation & Measurement Magazine, IEEE. – 2001. Т. 4. №. 2. P. 22–27.
3. *Franke U., Kutzbach I.* Fast stereo based object detection for stop&go traffic //Intelligent vehicles symposium, 1996., proceedings of the 1996 IEEE. – IEEE, 1996. – P. 339–344.
4. *Hirschmüller H.* Stereo processing by semiglobal matching and mutual information //Pattern Analysis and Machine Intelligence, IEEE Transactions on. – 2008. Т. 30. №. 2. P. 328–341.
5. *Hartley R. I.* Theory and practice of projective rectification //International Journal of Computer Vision. – 1999. – V. 35. – №. 2. P. 115-127.

# Исследование микротечений газа с помощью гибридного метода

О.И. Ровенская<sup>40</sup>

**Аннотация:** Гибридный метод, динамически сращивающий решение S-модельного кинетического уравнения с решением уравнений Навье-Стокса применяется для численного моделирования течения газа в микроканале, вызванного произвольным перепадом давления в широком диапазоне числа Кнудсена. Общее решение достигается сращиванием полу потоков макровеличин на границе между областями, обеспечивая сохранение массы, импульса и энергии через границу. Эффективность гибридного метода увеличена за счет распараллеливания с помощью MPI. Точность и надежность метода оценивается сравнением с полным решением кинетического уравнения.

**Ключевые слова:** микротечения; гибридный метод; кинетические уравнения

## Введение

Бурное развитие нанотехнологий потребовало разработки специальных подходов для их численных исследований. Несмотря на достигнутые успехи, микротечения являются объектом интенсивного исследования, а их изучение до сих пор представляет сложность [1-5]. Сосуществование разреженных и сплошносредних областей является характерной особенностью микротечений. Разреженные течения описываются с помощью кинетических подходов, которые являются численно затратными вследствие дискретизации кинетического уравнения в физическом и скоростном пространствах. Сплошносредные течения хорошо описываются менее затратными уравнениями Навье-Стокса, которые могут быть менее точными в критических разреженных областях. В пограничном слое неточность

---

<sup>40</sup> ФИЦ ИУ РАН (ВЦ им. А. А. Дородницына РАН); 1982 г.р.; к.ф.-м.н; olga\_rovenskaya@mail.ru

частично преодолевается введением условий скольжения и температурного скачка на твердой поверхности, которые справедливы при числе Кнудсена  $Kn \leq 0.1$ . Кроме того, в микро течениях разреженность играет ключевую роль и уравнения Навье-Стокса даже с учетом условий скольжения не всегда справедливы во всем поле течения.

Таким образом, создание гибридных методов, использующих кинетический подход в локальных разреженных областях течения и сплошносреднее описание в остальной части течения, для моделирования многомасштабных течений является важной областью исследования в последние годы. Наиболее популярным подходом при создании гибридного метода является разбиение физического пространства на кинетическую и сплошносреднюю подобласти, используя некоторый критерий. При этом в кинетической подобласти в основном используются методы Монте-Карло. В разработанном гибридном подходе [3-5] применяется прямое численное решение кинетического уравнения, что предпочтительнее для процедуры сращивания решений, т.к. в этом случае и кинетический, и гидродинамический подходы используют схожие численные схемы. Для расщепления течения на подобласти используется критерий, основанный на градиентах макровеличин и локальном  $Kn$ , предложенный в [6].

Гибридный метод используется для исследования течения разреженного газа в плоском микроканале при произвольном перепаде давления в нем. Рассматривается предельный случай – течение через щель, когда эффектом твердых стен можно пренебречь.

## 1. Постановка задачи

Рассматривается двумерное течение монокатомного газа в канале высоты  $H$  и длины  $l = 10H$ , соединяющего резервуары размера  $L_x \times L_y = 5H \times 3H$  (см. Рис. 1), заполненных газом. В первом резервуаре задается полная температура  $T_0$ , давление  $p_0$  и направление потока. Статическое давление  $p_e$  задается во втором.

$$\frac{p_0}{p_e} = \left(1 + \frac{\gamma-1}{2} Ma_{is}^2\right)^{\gamma/(\gamma-1)} \quad \frac{T_0}{T_e} = \left(1 + \frac{\gamma-1}{2} Ma_{is}^2\right) \quad (1)$$

здесь  $Ma_{is}$  изоэнтропическое число Маха,  $\gamma = c_p/c_v$  отношение теплоемкостей,  $p_0 = R\rho_0 T_0$ ,  $R$  газовая постоянная,  $\rho$  плотность газа. Температура стен  $T_w$  совпадает с температурой газа  $T_0$ . В случае  $I = 0$  газ течет через щель. Течение симметрично относительно  $y = 0$ , поэтому только нижняя половина области моделируется. Течение газа определяется величиной перепада давления  $p_0/p_e$  и параметром разреженности газа  $\delta = p_0 H/\mu_0 v_0$ , здесь  $v_0 = (2RT_0)^{0.5}$  тепловая скорость,  $\mu_0$  динамическая вязкость газ при  $T_0$ .

Безразмерный поток массы  $W = m/m_{fm}$  является глобальной характеристикой течения, и нормируется на свободномолекулярное значение  $m_{fm} = p_0/\pi^{0.5} v_0$ . Поток массы через канал  $m$  равен:

$$\dot{m} = 2 \int_0^{H/2} \rho(x, y) u(x, y) dy \quad (3)$$

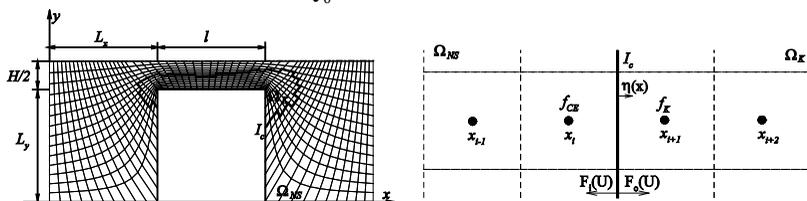


Рис. 1. Область течения и схема срачивания решений.

## 2. Гибридный метод

Схема срачивания решений показана на Рис. 1. Для определения размера и положения кинетической подобласти  $\Omega_K$  выбран критерий, основанный на оценке локального числа Кнудсена  $Kn$  и градиентов макропараметров  $Kn_{GL}(\mathbf{x})$  [3-6]:

$$Kn_{GL}(\mathbf{x}) = \max(Kn_{GLp}, Kn_{GLV}, Kn_{GLT}) \quad Kn_{GL\Phi}(\mathbf{x}) = Kn |\nabla\Phi|/\Phi \quad (4)$$

здесь  $\Phi = (\rho, |\mathbf{V}|, T)$  вектор макропараметров течения. Если  $K_{\text{LGL}}(\mathbf{x}) \geq \Xi$  решается кинетическое уравнение, в остальной области решаются уравнения Навье-Стокса.  $\Xi = 0.1$  гарантирует разность между гибридным и кинетическим решением около 1% [5]. Если  $\mathbf{x} \in \Omega_K$  решается S модельное уравнение [7]

$$\frac{\partial f}{\partial t} + \boldsymbol{\xi} \frac{\partial f}{\partial \mathbf{x}} = \delta \rho \sqrt{T} (f_s - f) \quad (5)$$

здесь  $f = f(t, \mathbf{x}, \boldsymbol{\xi})$  функция распределения молекул по скоростям, т.е. вероятность обнаружить частицу со скоростью  $\boldsymbol{\xi} = (\xi_x, \xi_y, \xi_z)$  в точке пространства  $\mathbf{x} = (x, y)$  в момент времени  $t$ .  $f_s$  стандартная локальная функция Шахова [7],  $\mathbf{c} = \boldsymbol{\xi} - \mathbf{V}$  собственная скорость молекул,  $\mathbf{V} = (u, v)$  массовая скорость газа. Для обезразмеривания (5) использовались равновесные параметры газа  $\rho_0$  и  $T_0$  и высота микроканала  $H$ . Применяется модель твердых шаров, тогда безразмерная вязкость  $\mu = T^{0.5}$ .

В остальной, сплошнородной области течения  $\Omega_{NS}$  решаются вязкие сжимаемые уравнения Навье-Стокса:

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{U})}{\partial \mathbf{x}} = 0 \quad (6)$$

здесь  $e_{\text{tot}} = e + \mathbf{V}^2/2$  полная энергия в единице массы,  $\mathbf{U} = (\rho, \rho \mathbf{V}, \rho e_{\text{tot}})$  вектор макровеличин,  $\mathbf{F}(\mathbf{U})$  вектор потоков, включающий конвективные и диффузионные (вязкие) компоненты. В кинетическом описании макровеличины плотность  $\rho$ , импульс  $\rho \mathbf{V}$  и внутренняя энергия единицы массы  $e$  вычисляются интегрированием по всему скоростному пространству.

Так как течение является двумерным можно аналитически упростить задачу, убрав компоненту вектора скорости  $\xi_z$  [1, 5]. Кинетическое уравнение (5) дискретизируется в

физическом и скоростном пространствах и решается, используя явно-неявную схему [1, 5]. Этап переноса в уравнении (5) аппроксимируется явной конечно-объемной схемой. Потоки вычисляются с помощью TVD схемы с ограничителем в виде  $\min\text{mod}$ .

Численное решение уравнение Навье-Стокса основано на конечно-разностном и конечно-объемном методах второго порядка точности по пространству и времени [5]. Для интегрирования по времени используется схема Кранка-Николса.

Сращивание между решениями уравнений Навье-Стокса и кинетического уравнения достигается согласованием полупотоков массы, импульса и энергии через границу между областями  $\Omega_c$  [5]. На границе  $\Omega_c$  для молекул входящих в  $\Omega_K$  задается функция распределения по скоростям Чепмена-Энскога с макропараметрами, вычисленными в  $\Omega_{NS}$ . Полупотоки массы, импульса и энергии входящие в подобласть  $\Omega_{NS}$  из  $\Omega_K$  вычисляются на основе решения кинетического уравнения в  $\Omega_K$  (см. Рис. 1). При этом автоматически выполняется сохранение потока массы, импульса и энергии через границу  $\Omega_c$ . Положение границы  $\Omega_c$  между  $\Omega_K$  и  $\Omega_{NS}$  определяется на каждом шаге по времени. Если точка пространства находящаяся в  $\Omega_{NS}$  на предыдущем шаге по времени перемещается в  $\Omega_K$ , то функция распределения в ней задается в виде функции Чепмена-Энскога.

Граничные условия в  $\Omega_{NS}$  задаются с помощью условий (1). Если входная/выходная границы находятся в  $\Omega_K$ , для молекул, входящих в область, задается локально-максвелловская функция распределения с параметрами, соответствующими условиям в резервуарах. На твердой поверхности в  $\Omega_K$  задаются диффузные граничные условия с полной аккомодацией. Если твердая поверхность попадает в  $\Omega_{NS}$ , то задаются условия скольжения и температурный скачек.

Для повышения эффективности гибридный код распараллелен с помощью MPI (message passing protocol).

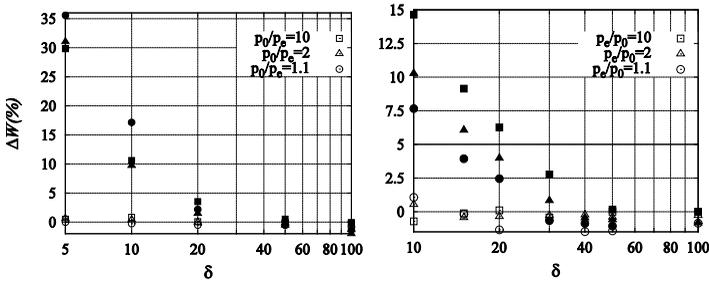


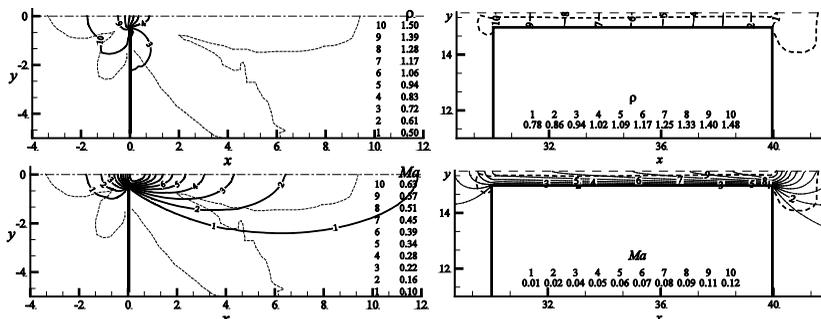
Рис. 2.  $\Delta W_{NS}$  (сплошные символы) и  $\Delta W_h$  (пустые символы): щель (слева), канал (справа)

### 3. Анализ численных результатов

Оценим отклонение решений полученных гибридным методом  $\Delta W_h$  и с помощью решения уравнений Навье-Стокса  $\Delta W_{NS}$  от кинетического решения как  $W_{(h, NS)} = 1 - W_{(h, NS)}/W_s$ . На Рис. 2 показаны относительные разности  $\Delta W_{NS}$  и  $\Delta W_h$  для течения через щель и канал при  $p/p_e = 10, 2$  и  $1.1$  и параметре разреженности  $\delta$  от 5 до 100. Когда газ слабо разрежен ( $\delta \geq 50$ ) потоки массы, полученные кинетическим, гибридным и Навье-Стокс методами, отличаются не более чем на 2%. Начиная с  $\delta = 20$  разница между кинетическим и решением уравнений Навье-Стокс  $\Delta W_{NS}$  становится заметной ( $>5\%$ ) и растет с увеличением разреженности газа. Следует отметить, что в случае течения через щель, превышение 5% порога достигается позже, при  $\delta = 15$ , чем для течения в канале, вследствие слабого воздействия твердых границ на течение через щель. Максимальное отклонение  $\Delta W_{NS} = 35\%$  для течения через щель при  $\delta = 5$  и  $15\%$  при  $\delta = 10$  для течения в канале. При этом разность между гибридным и кинетическим решением  $\Delta W_h$  порядка 1-2%, во всем рассматриваемом диапазоне параметра  $\delta$ .

Распределение контурных линий плотности и числа Маха около щели и вдоль канала, а также граница  $I_c$  между  $\Omega_K$  (внутри) и  $\Omega_{NS}$  показаны на Рис. 3 при  $\delta = 10$  и  $p/p_e = 2$ . Локальное число  $Kn$  около  $I_c$  изменяется от 0.067 до 0.4 для

щели и от 0.092 до 0.17 в канале, при этом максимальный  $Kl$  возникает в окрестности щели и равен 1.53, а на выходе из канала равен 0.19. Можно видеть, что линии, пересекающие границу  $I_c$ , демонстрируют плавный переход, что говорит о правильном сращивании кинетического и Навье-Стоксовского решений.



**Рис. 3.** Плотность и число Маха:  $\delta = 10$  и  $\rho_0/p_e = 2$ .

Время, требуемое для одного шага гибридного кода,  $t_h$  представляет собой сумму времен, требующихся для решения S-модельного уравнения в  $N_K$  кинетических точках, уравнений Навье-Стокса в  $N$   $N_K$  точках и процедуры сращивания. Так как время, затрачиваемое на решение уравнения Навье-Стокса и процедуру сращивания достаточно мало, то  $t_h$  определяется числом  $N_K$  точек, в которых решается модельное уравнение. Достигнутое гибридным методом ускорение оценивалось как  $s = t_K/t_h$ , где  $t_K$  время решения кинетического уравнения в  $N$  точках. Максимальное ускорение, достигаемое в расчетах  $s = 12$  при  $\delta = 100$ , и падает с уменьшением  $\delta$ , до 2 при  $\delta = 5$ , т.к. с увеличением разреженности газа растет число кинетических точек  $N_K$ .

## Выводы

Гибридный метод, основанный на сращивании решений кинетического уравнения и уравнений Навье-Стокса использовался для исследования течения в микроканале и

щели. Показано, что результаты, полученные гибридным методом, совпадают с кинетическими, даже при тех условиях, когда уравнения Навье-Стокса дают полностью некорректные результаты. При этом гибридный метод демонстрирует ускорение (от 12 до 2) в сравнении с решением кинетического уравнения.

## Список литературы

1. *Rovenskaya O.* Comparative analysis of the numerical solution of full Boltzmann and BGK model equations for the Poiseuille flow in a planar microchannel // *Computers&Fluids*, 2013. Vol. 81.P. 45–56.
2. *Rovenskaya O. I.* Kinetic analysis of surface roughness in a microchannel // *Computers&Fluids*, 2013. Vol. 77.P. 159–165.
3. *Rovenskaya O., Croce G.* Coupling kinetic and continuum equations for micro scale flow computations // *Heat Transfer Engineering*, 2013. Vol.34. No. 2-3. P. 192–203.
4. *Rovenskaya O., Croce G.* Numerical investigation of micro-flow over rough surfaces: coupling approach // *Journal of Heat Transfer*, 2013. Vol. 135. No 10. P. 101005.
5. *Rovenskaya O., Croce G.* Application a hybrid solver to gas flow through a slit at arbitrary pressure ratio // *Vacuum*, 2014. Vol. 109. P. 266–274.
6. *Schwartzentruber T. E., Boyd I. D.* A hybrid particle-continuum method applied to shock waves // *Journal of Computational Physics*, 2006. Vol. 215. P. 402–416.
7. *Шахов Е. М.* Метод исследования движения разреженного газа. – М.: Наука, 1974.

# Методы и средства информатики для диагностики артериальной гипертензии в лечебно-профилактических учреждениях широкого профиля

С. Б. Румовская<sup>41</sup>

**Аннотация:** Представления об организме человека и процессе его диагностики как гетерогенной системе еще не завершены в теории и не стали картиной мира клинической врачебной практики. Целостное восприятие человека в клинической медицине остаётся феноменом высочайшего искусства. В условиях существенного разнообразия диагностической информации актуален переход от устоявшегося клинического гомогенного подхода в диагностике к новому гетерогенному подходу. В связи с чем предлагается улучшить качество медицинских решений за счет консультаций у искусственного коллективного интеллекта, имитирующего в компьютере коллективную работу врачей при диагностике пациента. Предлагается инструментальная среда «Виртуальный консилиум» для поддержки принятия решений в процессе диагностики артериальной гипертензии.

**Ключевые слова:** поддержка принятия диагностических решений; артериальная гипертензия; виртуальный консилиум

## Введение

Представления об организме человека и процессе его диагностики как гетерогенной системе еще не завершены в теории и не стали картиной мира клинической врачебной практики. Несмотря на стремление биомедицины к гетерогенности восприятия организма человека и процесса его диагностики, человек остается «расчлененным», редуцированным по дисциплинарным «полочкам» объектом познания. Исследования

---

<sup>41</sup> ФИЦ ИУ РАН (Калининградский филиал); 1985 г.р.; [sophiyabr@gmail.com](mailto:sophiyabr@gmail.com)

внутриклеточных процессов сформировали специалистов узкого профиля, поглощенных решением частных задач. Новый тип ученого «прагматика-фактолога» перестал задумываться над тем, что делается «вокруг» и какое значение могут иметь добытые им факты для понимания работы организма в целом.

В условиях существенного разнообразия и объемов перерабатываемой диагностической информации актуален переход к гетерогенному подходу в диагностике. В связи с чем предлагается поднять качество медицинских решений в лечебно-профилактических учреждениях (ЛПУ) широкого профиля за счет консультаций с моделью [1], имитирующей в компьютере коллективную работу врачей при диагностике пациента, в частности, диагностике артериальной гипертензии (АГ). В данной работе представлены результаты исследования диагностического процесса в стационаре и ЛПУ широкого профиля и метод применения инструментальной среды «Виртуальный консилиум» для поддержки принятия решений в процессе диагностики АГ.

## **1. Диагностика артериальной гипертензии в стационаре больницы**

Анализ проблемной среды диагностики АГ [2] на примере Калининградской клинической областной больницы (стационара, рис. 1) показал, что для формирования полного дифференциального диагноза артериальной гипертензии привлекаются до тринадцати врачей-экспертов (на рис. 1 –  $M = \overline{1,13}$ ) – носителей знаний из различных разделов медицины: невролог, нефролог, сосудистый хирург, уролог, психолог, педиатр, акушер-гинеколог, онколог, окулист, врачи функциональной диагностики, эндокринолог, терапевт, кардиолог – лицо, принимающее решение (ЛПР). Они сотрудничают и дополняют друг друга, рассматривая организм пациента с различных точек зрения, каждый относительно своего опыта и теоретической подготовки.



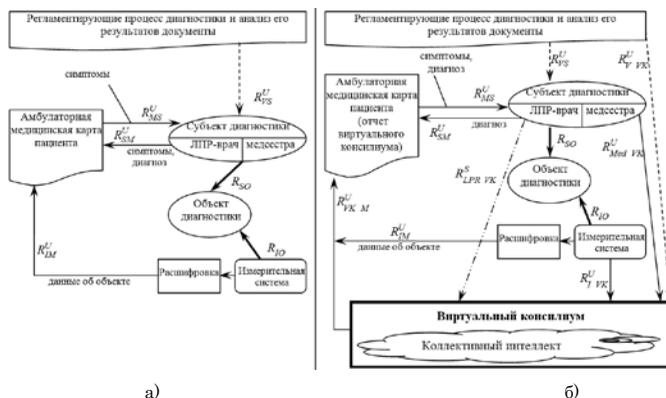
измерительной системой после расшифровки врачами функциональной диагностики (на рис. 1,  $R_{TM}^U$ ). Полученные исходные данные об объекте диагностики разнородны: измерения дают количественные и визуально-графическом параметры (детерминированные переменные). От ЛПР и врачей-экспертов данные поступают в формате количественных, лингвистических четких и нечетких переменных. Исследования концептуальной модели с позиции диагностики АГ показали большую размерность, порядка 800 значений показателей состояния здоровья пациента.

ЛПР изучает (на рис. 1,  $R_{MS}^U$ ) в медицинской карте симптомы и частные диагностические мнения врачей-экспертов, множество которых подбирает сам ( $R^S$ ), и ставит заключительный диагноз ( $R_{SM}^U$ ). Врачам-экспертам также доступны симптомы и мнения других врачей-экспертов в медицинской карте ( $R_{ME}^U$ ). ЛПР и врачи-эксперты обследуют пациента и формируют диагностические заключения по нормативным документам, например [3] ( $R_{VE}^U$ ,  $R_{VS}^U$ ).

## **2. Диагностика артериальной гипертензии в лечебно-профилактических учреждениях широкого профиля и «Виртуальный консилиум»**

Проблемная среда диагностики артериальной гипертензии [2] характеризуется существенным разнообразием, как в охватываемых разделах медицины, так и в типах симптоматической информации, её растущими объемами, инертностью врачей из-за сомнения в степени риска, связанного с высоким артериальным давлением, и инертностью при нормальном артериальном давлении на первом приеме. Как следствие – у 70% пациентов ЛПУ широкого профиля диагноз АГ «недовыявлен» и в основном из-за инертности. В то время как в стационарных медицинских учреждениях (рис. 1) практически в 100% имеет

место, как адекватное проведение, так и отображение в медицинских картах данных анамнеза и физикального обследования с подтверждением диагноза лабораторно-инструментальными методами исследования. В этой связи предлагается распространить опыт и знания стационарной диагностики АГ в ЛПУ широкого профиля (рис. 2а) за счет внедрения информационной технологии «Виртуальный консилиум» (рис. 2б).



Обозначения: а):  $\dashrightarrow$  – информационно-управляющие отношения «внешняя среда (регламентирующие документы) – субъект диагностики»;  $\longrightarrow$  – информационно-управляющие отношения «субъект диагностики – медицинская карта», «медицинская карта – субъект диагностики», «измерительная система – медицинская карта»;  $\Longrightarrow$  – отношения формирования множества значений показателей состояния здоровья объекта диагностики;

б):  $\dashrightarrow$  – информационно-управляющие отношения «внешняя среда – субъект диагностики» и «внешняя среда – виртуальный консилиум»;  $\dashrightarrow$  – отношения сотрудничества «субъект диагностики – виртуальный консилиум»;  $\longrightarrow$  – информационно-управляющие отношения «виртуальный консилиум – медицинская карта», «субъект диагностики – медицинская карта», «медицинская карта – субъект диагностики», «измерительная система – медицинская карта», «измерительная система – виртуальный консилиум»;  $\Longrightarrow$  – отношения формирования множества значений показателей состояния здоровья объекта диагностики;

**Рис. 2.** Концептуальная модель диагностического процесса в лечебно-профилактическом учреждении широкого профиля: а) без «Виртуального консилиума»; б) с применением «Виртуального консилиума»

В ЛПУ широкого профиля (рис. 2а) субъект диагностики – это врач общей практики или терапевт (иногда кардиолог, но

зачастую без опыта работы, к которому направляет терапевт сразу же при выявлении повышенного артериального давления), это врач «прагматик-фактолог», объединяющий в себе роли ЛПР-врача и экспертов-врачей узкой специализации. Субъект диагностики включает и медсестру, помогающую врачу на приеме пациентов.

Предлагаемая технология «Виртуальный консилиум» [1] и первая версия инструментальной среды для поддержки принятия диагностических решений по АГ, архитектура которой описана в [4], а структура в [1], ограничена работой со знаниями шести врачей-экспертов разной специализации, решающими двенадцать функциональных подзадач, возникающих в 90% случаев диагностики АГ, и также ограничивается пациентами старше 18, без особых состояний, нет распознавания снимков, не предусматривается назначение лечения и не диагностируется ряд симптоматических артериальных гипертензий.

При лабораторной апробации [1] прототипа «Виртуального консилиума» вычислялась среднеквадратическая ошибка постановки диагноза. В итоге диагноз прототипа верен в 84% случаев. Кроме этого, исследовалось время диагностики. Оно складывается из трех составляющих: времени на собственно диагностику и занесение информации в карту, времени обработки информации и времени записи заключения. Сумму последних двух составляющих с применением прототипа виртуального консилиума сократилось, примерно, до 30 сек.

Предварительные результаты лабораторных экспериментов с прототипом дали обнадеживающие результаты, и позволяют предположить, что внедрение «Виртуального консилиума» в ЛПУ широкого профиля повысит эффективность и качество индивидуальных диагностических решений ЛПР-врача до уровня стационара, поскольку консультации будут получены у искусственного коллективного интеллекта, обладающего синергией. При этом ЛПР-врач (рис. 2б) обследует пациента, а медсестра заносит симптомы в компьютер. «Виртуальный консилиум» анализирует полученную информацию и выводит за-

ключительный диагноз в форме аналогичной выписке из истории болезни стационара. ЛПР-врач при несогласии с диагнозом может повторно обратиться к «Виртуальному консилиуму».

## **Выводы**

Лабораторно подтверждена эффективность предлагаемого средства поддержки принятия решений при диагностике артериальной гипертензии и метода его применения. Предварительные результаты экспериментов с исследовательским прототипом позволяют предположить, что «Виртуальный консилиум», доведенный до промышленного прототипа, может улучшить качество диагностики артериальной гипертензии в лечебно-профилактических учреждениях широкого профиля.

## **Список литературы**

1. *Кириков И.А., Колесников А.В., Румовская С.Б.* Исследование лабораторного прототипа искусственной гетерогенной системы для диагностики артериальной гипертензии // Системы и средства информатики, 2014. Т.24. No.3. С. 121–144.
2. *Кириков И.А., Колесников А.В., Румовская С.Б.* Исследование сложной задачи диагностики артериальной гипертензии в методологии искусственных гетерогенных систем // Научный журнал «Системы и средства информатики», 2013. Т. 23. No. 2. С. 96–114.
3. *ESH/ESC.* 2013. Guidelines for the management of arterial hypertension: The Task Force for the management of arterial hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC) / Journal of Hypertension. 31(7): 1281–1357.
4. *Кириков И.А., Колесников А.В., Румовская С.Б.* Функциональная гибридная интеллектуальная система для поддержки принятия решений при диагностике артериальной гипертензии // Системы и средства информатики, 2014. Т.24. No.1. С. 153–179.

# Модель бизнес-процесса в виде сети массового обслуживания

О.В. Семенова<sup>42</sup>, А.С. Власкина<sup>43</sup>, М.А. Норкина<sup>44</sup>,  
К.А. Шевякова<sup>45</sup>

**Аннотация:** В работе на примере бизнес-процесса получения кредита производится расчет и оценка среднего времени пребывания заявки в системе. Для расчета и оценки данной характеристики использованы методы теории массового обслуживания и имитационного моделирования.

**Ключевые слова:** BPMN, бизнес-процесс, среднее время пребывания в системе, сеть массового обслуживания, имитационного моделирование, экспоненциальная одноуровневая сеть.

## Введение

Одним из основных показателей эффективности работы любой компании является время, которое она тратит на выполнение какого-либо отдельно взятого бизнес-процесса. Для оценки данной характеристики необходимо провести качественный и количественный анализ процесса [1,2]. В настоящее время для разработки и оптимизации бизнес-процессов чаще всего используются методы математического [3] и имитационного моделирования [4]. Данные методы позволяют оценить эффективность существующих в компании бизнес-процессов, а также определить недостатки в текущей организации процесса.

В качестве примера применения методов математического и имитационного моделирования рассмотрим анализ среднего времени выполнения бизнес-процесса.

---

<sup>42</sup> РУДН; 1994 г.р.; oksana.s.biv@mail.ru

<sup>43</sup> РУДН; 1995 г.р.; vlaskina.anastasia@yandex.ru

<sup>44</sup> РУДН; 1995 г.р.; marynor@mail.ru

<sup>45</sup> РУДН; 1993 г.р.; ksenia-sh@mail.ru

## 1. Описание бизнес-процесса

Для расчета среднего времени пребывания заявки в сети массового обслуживания (СеМО) в данной работе используются два метода: метод расчета среднего времени пребывания заявки в СеМО для экспоненциальной однородной сети и метод имитационного моделирования. Описанные методы реализованы для бизнес-процесса получения кредита, BPMN (Business Process Model and Notation) модель которого изображена на Рис. 1 [5].

Данная модель описывает процесс оформления заявки на кредит. В систему поступает заявка. Она попадает в очередь, после чего работники проверяют кредитную историю и источники дохода клиента. Затем кредитор проводит оценку заявки и принимает решение о выдаче кредита. В случае положительного ответа составляется кредитное предложение, и заявка считается обработанной. В случае отрицательного ответа кредитор уведомляет клиента об отказе, связываясь с ним через систему. В этой ситуации клиент может повторно подать запрос или отказаться от предоставления услуги. Вне зависимости от решения, принятого клиентом, заявка будет считаться обработанной. На этом процесс завершается.

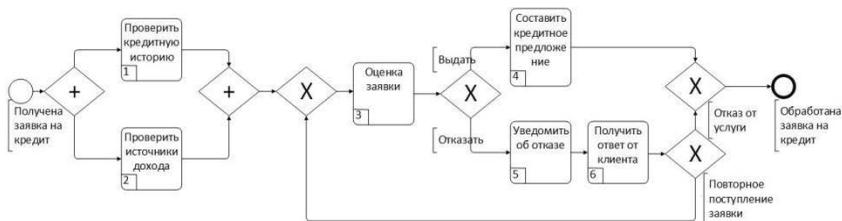
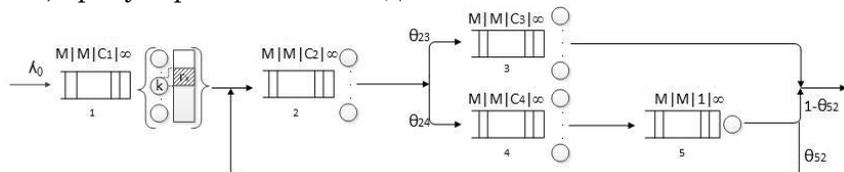


Рис. 1. BPMN модель

## 2. Среднее время пребывания заявки в СеМО

Для расчета среднего времени пребывания заявки в

СеМО (1) построена математическая модель, состоящая из 5 узлов, пронумерованных последовательно в соответствии с Рис. 2.



**Рис. 2.** Модель однородной экспоненциальной СеМО

В данной модели запросы клиентов на получение кредита соответствуют заявкам, а сотрудники и системы компании – узлам открытой однородной СеМО. Обозначим множество узлов за  $M=\{1,2,3,4,5\}$ ,  $|M|=5$ .

$\Delta$

$$= \left( \bar{\omega}_1 + \frac{1}{\mu_1} \right) + \frac{\left[ \left( \bar{\omega}_2 + \frac{1}{\mu_2} \right) + \theta_{23} \cdot \left( \bar{\omega}_3 + \frac{1}{\mu_3} \right) + \theta_{24} \cdot \left( \bar{\omega}_4 + \frac{1}{\mu_4} + \frac{1}{\mu_5 - \lambda} \right) \right]}{1 - \theta_{52}} \quad (1)$$

$\bar{\omega}_i$  – среднее время ожидания начала обслуживания,  $i = \overline{1,5}$ .

Для расчета среднего времени установления соединения построена математическая модель в виде открытой экспоненциальной СеМО. Узлы СеМО пронумерованы в прямом направлении (от узла, соответствующего М1, до узла М5). Ниже в таблице 1 перечислены классы заявок.

**Таблица 1.** Классы заявок

Заявки	Значение
1	Необработанные заявки
2	Обработанные заявки с положительным результатом
3	Обработанные заявки с отрицательным результатом (отказ клиента и/или банка)

Для данной модели множества заявок для каждого из классов будет выглядеть следующим образом:

$$L_1 = \{(1,1), (2,1)\},$$

$$L_2 = \{(3,2), (4,2), (5,2)\},$$

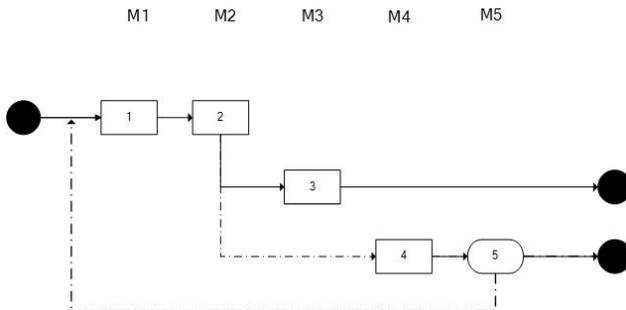
$$L_3 = \{(3,3), (4,3), (5,3)\}.$$

После окончания обслуживания в узле  $i$  заявка  $(i, r)$  становится  $(j, s)$ - заявкой, с вероятностью  $\theta_{(i,r),(j,s)}$ ,  $i, j \in M$ ,  $r, s \in R$ . Тогда  $\Theta = (\theta_{(i,r),(j,s)})_{i,j \in M, r,s \in R}$  - матрица переходных вероятностей цепи Маркова на множестве состояний  $L_r$ ,  $r \in R$ , как представлено в Таблице 2.

**Таблица 2.** Матрица переходных вероятностей

$\Theta$	0	1	2	3	4	5	$E$
0	0	1	0	0	0	0	1
1	0	0	1	0	0	0	1
2	0	0	0	$\theta_{23}$	$\theta_{24}$	0	1
3	1	0	0	0	0	0	1
4	0	0	0	0	0	1	1
5	$1 - \theta_{52}$	0	$\theta_{52}$	0	0	0	1

На Рис. 3 представлены возможные переходы между состояниями. Переходы, которые соответствуют множествам  $L_1$  и  $L_2$ , показаны сплошными стрелками, а переходы, соответствующие множеству  $L_3$ , - пунктирными стрелками. Состояния, соответствующие узлам  $M | M | C | \infty$ , обозначены на Рис.3 прямоугольниками, а состояния, соответствующие узлам  $M | M | 1 | \infty$ , - прямоугольником с закругленными углами, узлы, соответствующие внешней среде, - черными кругами.



**Рис. 3.** Граф переходов между состояниями однородной СеМО

### 3. Пример численного анализа

В ходе имитационного моделирования были получены результаты, приведенные на Рис. 4 серой кривой. Черной кривой обозначены результаты подсчетов с помощью математической модели.

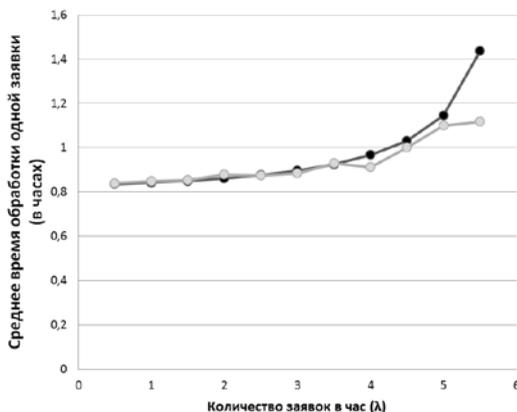


Рис. 4. Среднее время обслуживания заявки на кредит

### Выводы

В докладе приведены результаты анализа одного из наиболее важных показателей эффективности выполнения бизнес-процесса – среднего времени оформления заявки на кредит. Расчет проводился методами анализа СеМО и средствами имитационного моделирования.

### Список литературы

1. Самуйлов К.Е., Чукарин А.В., Яркина Н.В. Бизнес-процессы и информационные технологии в управлении телекоммуникационными компаниями. – М.: Альпина Пабlishерз, 2009. – 442 с.
2. Dumas M., La Rosa M., Mendling J., and Reijers H.

Fundamentals of business process management // Springer, 2013. – 399 P

3. *Башарин Г.П.* Лекции по математической теории телетрафика: Учеб. пособие. – Изд. 3-е, испр. и доп. – М.: РУДН, 2009. – 342 с
4. The Business Process Simulator (доступ: <http://bimp.cs.ut.ee/>).
5. *Гудкова И.А., Власкина А.С., Семенова О.В., Норкина М.А.* К анализу среднего времени выполнения бизнес-процесса методами теории массового обслуживания и имитационного моделирования // Всероссийская конференция с международным участием «Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем» ИТТММ-2015 (20–24 апреля 2015 г., Москва): Материалы. – М.: РУДН. – 2015. – С. 15–17.

# Реляционная модель хранения данных информационной поддержки задач структурного материаловедения\*

П. А. Сеченых<sup>46</sup>, К. К. Абгарян<sup>47</sup>

**Аннотация:** Работа посвящена вопросу разработки и реализации информационной поддержки задач структурного материаловедения. При моделировании многих процессов, проходящих в кристаллах, а также при расчете их свойств, часто возникает необходимость использовать различные подходы, такие как методы молекулярной динамики, квантово-механические методы и т.д., а также создавать на их основе соответствующие программные модули и многомасштабные схемы расчета [1]. Часто встает вопрос о передаче данных между расчетными модулями. Представлена реляционная модель хранения данных, применяемая для решения ресурсоемких и разномасштабных задач структурного материаловедения. Разработанная на основе этой модели база данных является центральным блоком многомасштабной схемы расчета: с ее помощью осуществляется обмен данными между различными вычислительными модулями.

**Ключевые слова:** база данных; реляционная модель хранения; многомасштабная схема расчета

## Введение

Задача получения новых материалов с заданными параметрами и прогнозируемыми свойствами является актуальной и весьма ресурсоемкой в силу ряда причин. В рассматриваемой предметной области физическая модель представляется на молекулярном или атомном уровне. Из-за этого возникает необходимость обрабатывать (а именно: хранить, выбирать, готовить для расчета) большой объем данных. Кроме того, здесь

---

\*Работа выполнена при поддержке РФФИ (грант №13-08-01335А)

<sup>46</sup> ФИЦ ИУ РАН (ВЦ им. А. А. Дородницына РАН), МАИ; 1991 г.р.; p-sechenyh@mail.ru

<sup>47</sup> ФИЦ ИУ РАН (ВЦ им. А. А. Дородницына РАН), МАИ; 1963 г.р.; к.ф.-м.н.; kristal83@mail.ru

имеет место большое количество высокоточных вычислений и расчетов, схемы которых чувствительны к входным данным.

Исходя из этого, возникает необходимость разработать информационную поддержку многомасштабной схемы расчета, которая должна позволять хранить большие объемы типизированных различных расчетных данных, а также осуществлять передачу данных между вычислительными модулями.

## 1. Анализ предметной области

На начальном этапе работы был проведен анализ предметной области – структурного материаловедения, которое включает в себя кристаллографию, физику твердого тела, неорганическую химию и т.д. Физические свойства кристаллов обуславливаются особенностью их строения. Она заключается в том, что частицы, составляющие кристалл (атомы, ионы), располагаются закономерно с учетом пространственной симметрии, образуя тем самым трехмерную периодическую решетку [2,3]. В ходе анализа были классифицированы и структурированы кристаллографические данные (Рис. 1).

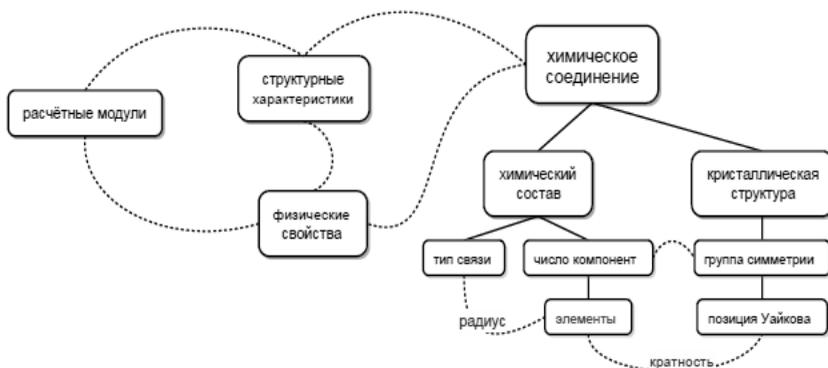


Рис. 1. Классификация данных предметной области

## 2. Модель хранения данных

По результатам проведенного анализа была разрабо-

тана схема базы данных для хранения соответствующей информации. Спроектированная схема базы данных в нотации «сущность-связь» [4] представлена ниже.

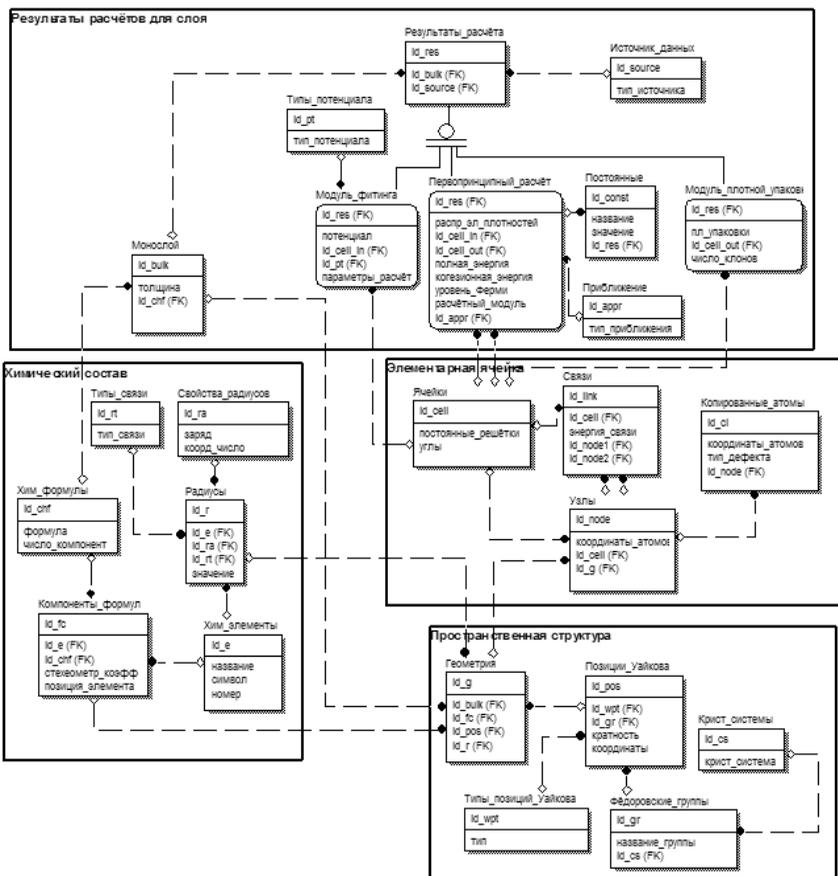


Рис. 2. Схема базы данных

В представленной на Рис. 2 структуре взаимосвязанных отношений можно выделить четыре подсхемы:

- *результаты расчетов*: каждому кристаллическому массиву, или монослою, ставятся в соответствие характеристики, полученные при помощи различных расчетных модулей (в

частности, модуль плотной упаковки, модуль квантово-механического расчета, молекулярно-динамический модуль), а также исходные данные расчета – химический состав и пространственная структура соединения.

- *данные о химическом составе:* химическая формула соединения и элементы таблицы Менделеева связаны отношением «многие-ко-многим», для его реализации использована ассоциативная сущность. Химические элементы и типы радиусов также связаны соотношением «многие-ко-многим» (поскольку каждому химическому элементу может соответствовать набор радиусов для различных типов связи). Для реализации данного соотношения вводится ассоциативная сущность, которая включает в себя значения радиусов.
- *данные о пространственной структуре соединения:* федоровская группа принадлежит к определенной кристаллической системе. Для каждой пространственной группы определен набор позиций Уайкова, ссылающихся на нее при помощи внешнего ключа. Позиция Уайкова принадлежит к определенному типу, который, в свою очередь, определяется соответствующей таблицей типов. Условие принадлежности атома какой-либо позиции Уайкова ограничивает область изменения его координат.
- *структурные характеристики элементарной ячейки:* элементарная ячейка представлена как собственными характеристиками (постоянные решетки и углы), так и набором составляющих ее атомов (хранятся рассчитанные координаты).

### 3. Реализация

Ниже представлены рисунки, иллюстрирующие работу приложения. На Рис. 3 показан пример заполнения формы. Рассматривается структура перовскита: это трехкомпонентное соединение, реализуемое в кубической группе симметрии  $R\bar{3}m$ , химическая формула соединения –  $\text{CaTiO}_3$ . В данном

примере в качестве критерия поиска пространственная группа не задана. На Рис. 4 показан результат выполнения поискового запроса по указанным параметрам: левая часть таблицы содержит информацию, общую для всего соединения (пространственная группа, постоянные решетки  $a$ ,  $b$ ,  $c$  и углы  $\alpha$ ,  $\beta$ ,  $\gamma$ , плотность упаковки), в правой части – характеристики атомов (химический элемент, позиция Уайкова, значение радиуса и координаты  $X$ ,  $Y$ ,  $Z$ ).

Редактор формул

число компонент: 3

формула: Ca 1 Ti 1 O 3

Группа симметрии: default

Фёдоровская группа:

позиции Уайкова (и их кратности):

Источник данных:

Модуль плотной упаковки

Модуль первопр. расчёта

поиск

Рис. 3. Пример задания входных данных

результаты поиска: Ca1Ti1O3

Ф.Г.	rho	a	b	c	alpha	beta	gamma	элемент	ПУ	радиус	X	Y	Z
Рез1м	0.627	4.02	4.02	4.02	90	90	90	O	a	148	0	0	0
								Ti	b	74.5	0.5	0.5	0.5
								O	c	126	0.5	0.5	0
Рез2м	0.246	5.48	5.48	5.48	90	90	90	Ca	a	148	0	0	0
								Ti	b	74.5	0.5	0.5	0.5
								O	d	126	0.5	0.5	0
Рез3м	0.299	10.2	10.2	10.2	90	90	90	Ca	g	148	0.313	0.313	0.313
								Ti	g	74.5	0.103	0.103	0.103
								O	k	126	0.35	0.174	0
Рез4м	0.295	10.3	10.3	10.3	90	90	90	Ca	g	148	0.356	0.356	0.356
								Ti	g	74.5	0.144	0.144	0.144
								O	f	126	0.5	0.322	0.314
Рез5м	3.92	3.92	3.92	90	90	90	90	Ca	a	148	0	0	0
								Ti	b	74.5	0.5	0.5	0.5
								O	c	126	0.5	0	0
								O	c	126	0	0.5	0
								O	c	126	0	0	0.5

сохранить

Рис. 4. Результат выполнения запроса

## Выводы

Проведен анализ предметной области, классифициро-

ваны и структурированы материаловедческие и кристаллографические данные. Построены реляционная модель хранения данных и правила агрегирования данных. Спроектирован пользовательский интерфейс, позволяющий осуществлять выборку данных согласно заданным критериям.

Реализованный прототип системы информационной поддержки апробирован на тестовых примерах (структура перовскита, полупроводники типа АПВВ).

Предложенный подход может быть использован также при расчете свойств многослойных полупроводниковых наноструктур (гетероструктур) [5]. В этом случае, расчетные данные по каждому слою хранятся отдельно, кроме того добавляются макросвойства всей гетероструктуры. Однако следует учесть, что такие материалы включают в себя промежуточные интерфейсные слои, которые обладают специфичным строением и свойствами.

## Список литературы

1. *Абгарян К.К.* Применение оптимизационных методов для моделирования многослойных полупроводниковых наносистем // Труды Института системного анализа Российской академии наук. Динамика неоднородных систем, 2010. Т.53(3).
2. *Шаскольская М.П.* Кристаллография. – М.: Высшая школа, 1976. – 390с. с ил.
3. *Hahn T.* International Tables for Crystallography. Vol. A. – 2005.
4. *Чен П.* Модель «сущность-связь» – шаг к единому представлению о данных // Системы управления базами данных, 1995.
5. *Абгарян К.К., Сеченых П.А., Гаврилов Е.С.* Объектно-реляционный подход к разработке системы компьютерного моделирования многомасштабной схемы расчета многослойных полупроводниковых наноструктур // Программная инженерия, 2015. №8. С 9–17.

# Метод детектирования машинопечатных текстовых информационных зон отпечатков документов строгого учета

И.И. Сидоркин<sup>48</sup>

**Аннотация:** Печатным документом признается отпечаток электронного документа, скрепленный собственноручной подписью и печатью владельца документа. Широко известны случаи злонамеренного внесения изменений в текстовую составляющую печатных документов, что может повлечь неверный ввод и дальнейшую обработку данных документов. В статье предлагается метод детектирования машинопечатных текстовых информационных зон отпечатков документов строгого учета для облегчения возможности распознавания текстового содержимого печатного документа и сравнения с электронной копией данного документа.

**Ключевые слова:** метод; цветной штриховой код; печатный документ; электронный документ; СВС; маркер; отпечаток документа; скан документа; машинопечатная текстовая информационная зона; алгоритм; документооборот

## Введение

Типовым регламентом внутренней организации федеральных органов исполнительной власти России в июле 2005 года утвержден порядок работы с электронными документами. В соответствии с ним, каждый электронный документ должен быть распечатан и только после этого обработан согласно утвержденной процедуре. При совместной работе над документами нескольких ведомств, обмен между ними ведется печатными документами на бумажных носителях. Данный регламент

---

<sup>48</sup> ФИЦ ИУ РАН (Орловский филиал), Приокский государственный университет; 1990 г.р.; voronecburgsiti@mail.ru

предусматривает использование системами электронного документооборота в качестве оригинала документа его печатного варианта [1].

Любой отпечаток печатного документа должен обладать свойствами уникальности и идентифицируемости. Данное свойство обеспечивается за счет нанесения немашиночитаемых данных: рукописный текст, оттиски печатей и штампов. В противном случае данный отпечаток не обладает требуемой юридической полнотой и не принимается в качестве документа строгого учета современными системами документооборота.

Несмотря на высокий уровень развития современных систем документооборота, все еще остро стоит проблема учета, регистрации и персонификации документов. Данные проблемы обусловлены необходимостью сканирования печатных документов, выделения отдельных информационных областей документов, распознавания, фильтрации и сохранения данных, и последующего перенаправления на дальнейшую обработку.

## **1. Постановка задачи**

Наделение бумажных отпечатков документов цветными машиночитаемыми объектами в системах электронного документооборота позволяет осуществить более удобный контроль над движением бумажных документов за счет значительного сокращения времени на поиск электронных копий документов при вводе и обработке их бумажных копий [1-7].

Информационная технология передачи данных на основе цветной печати, обеспечивающая реализацию документооборота средствами установления подлинности отпечатков электронных документов, реализована в [1-3,6] с помощью декодирования машиночитаемых объектов – цветных штриховых кодов СВС (Color Bar Code), содержащих ключевые данные электронного документа в закодированном виде.

Несмотря на все преимущества данной технологии, она не позволяет обрабатывать текстовую информацию, содержа-

щуюся в документе, которая доступна для прочтения держателем данного документа. Тем временем широко известны частые случаи злонамеренного внесения изменений в текстовую составляющую печатных документов, факт наличия которых, зачастую, может быть определен лишь путем проведения технической экспертизы документов специальными органами.

Осуществление контроля над достоверностью текстового содержимого печатного документа, его электронному образцу, возможно за счет взаимно-однозначного сравнения информации, содержащейся в печатном документе, и информации, хранимой в электронном документе.

Ввиду наличия большой вариативности форм документации строгого учета и искажений, получаемых при сканировании, встает проблема выделения и соотнесения для последующего сравнения отдельных участков печатного документа их электронному образцу.

подавляющее большинство документов строгого учета имеют сложную структуру, включающую в себя: таблицы, рамки, рукописный текст, печати и штампы, штриховой код, печатный текст различного начертания и т.п. Часть областей печатного документа не подлежит сравнению с электронным вариантом данного документа, поэтому она должна быть исключена из перечня областей на распознавание.

В связи с этим, выделение машинопечатных текстовых информационных зон в документе является актуальной задачей.

## **2. Описание метода**

В качестве решения описанной проблемы предлагается прибегнуть к использованию шаблонов документов (макет документа) для позиционирования отдельных областей информационного содержимого как электронного, так и печатного документа. Такой шаблон должен содержать информацию о координатах начала и конца каждой информационной области доку-

мента, которые необходимы для верификации. Большая вариативность форм строгой документации при этом требует наличия соответствующего количества шаблонов обрабатываемых документов.

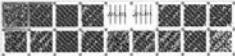
К основным ошибкам, возникающим в результате ввода печатной документации, можно отнести неточное позиционирование отпечатка документа на стекле сканера, что затруднит или полностью сделает невозможным процесс сегментации скана документа на отдельные информационные блоки по шаблону и сравнение их информационного содержимого с данными из базы.

Для разрешения данной проблемы предлагается использовать дополнительное информационное насыщение бумажных отпечатков документов маркерами, позволяющими однозначно позиционировать область документа для последующей обработки с помощью шаблона документа. При этом каждый шаблон документа дополняется информацией о местоположении маркеров. Изображение маркера представлено на Рис. 1.



**Рис. 1.** Маркер документа

Для осуществления точного позиционирования шаблона документа относительно скана печатного документа необходимо наличие минимум пары маркеров. Данное ограничение обусловлено необходимостью позиционирования левого верхнего и правого нижнего угла текстовой информационной зоны скана отпечатка документа. При этом ориентация документа (верх/низ) определяется за счет заведомо известного местоположения СВС относительно текстового содержимого конкретного документа. Тогда как угол искажения поворота определяется на основе информации о координатах найденных маркеров на скане и информации о координатах маркера в шаблоне документа. На Рис. 2 представлен скан платежного поручения с СВС и маркерами по углам текстовой информационной зоны скана отпечатка документа.



24.09.2015 24.09.2015  
 Поступ в банк плат. Слитоно о в. плат.

0410090

24.09.2015 Электронно  
 Дата Вид платежа 02

**ПЛАТЕЖНОЕ ПОРУЧЕНИЕ № 481**

Сумма прописью: шестьдесят восемь тысяч шестьсот семьдесят восемь, рублей 00 копеек

ИНН 755005550	КПП 75501001	Сумма	68678-00
ООО "Рассвет"		Сч. №	40703920007330010255
Плататель		БИК	047501711
ОАО "ЧЕЛИНЦБАНК" Г. ЧЕЛЯБИНСК		Сч. №	30101910500000000811
Банк плательщика		БИК	047501001
ГРКЦ ГУ БАНКА РОССИИ ПО ЧЕЛЯБИНСКОЙ ОБЛ. Г. ЧЕЛЯБИНСК		Сч. №	
Банк получателя		Сч. №	40101810400000010801
ИНН 7459000015	КПП 745901001	Вид оп.	01
УФК (Межрайонная ИФНС России №20 по Челябинской обл.)		Наз. пл.	
		Код	5
		Срок плат.	
		Смер. плат.	
		Рез. поле	
18210102010011000110	75626101001	ТП	МС.09.2015
			0 0 0

УИИД/И/валог на доходы физических лиц за 24.09.2015г.  
 Сумма 68678-00

Назначение платежа

Подпись \_\_\_\_\_

Отметка банка \_\_\_\_\_

М.П. \_\_\_\_\_

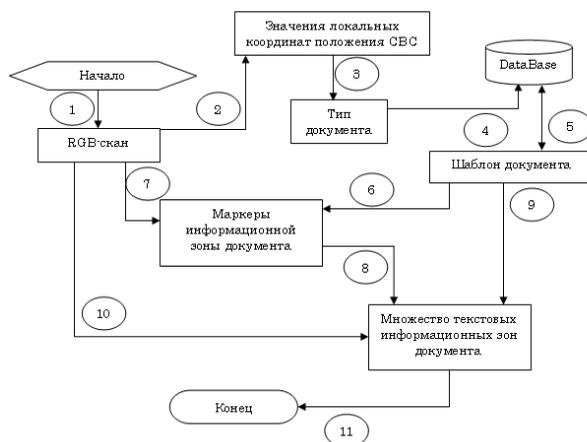
**Рис. 2.** Скан платежного поручения с СВС и маркерами

Алгоритм работы метода детектирования машинопечатных текстовых информационных зон документов строгого учета представлен на Рис. 3.

Процессами данного алгоритма являются:

- 1 – получение RGB-скана отпечатка документа;
- 2 – детектирование СВС на скане документа;
- 3 – декодирование СВС, получение данных о типе документа;
- 4 – формирование запроса к базе данных для поиска шаблона документа;
- 5 – обращение к

базе данных, получение шаблона документа; 6 – получение информации о местоположении маркеров на документе по данным из шаблона; 7 – детектирование маркеров информационной зоны документа на RGB-скане; 8 – совмещение найденных маркеров и маркеров шаблона документа; 9 – получение информации о машинопечатных текстовых информационных зонах документа из шаблона; 10 – получение координат машинопечатных текстовых информационных зон документа на RGB-скане; 11 – завершение работы процедуры.



**Рис. 3.** Алгоритм работы метода детектирования машинопечатных текстовых информационных зон документов строгого учета

## Выводы

Применение описанного в статье метода детектирования машинопечатных текстовых информационных зон документов строгого учета позволяет решить следующие задачи:

1. выделение информационной зоны скана документа;
2. согласование координат маркеров информационной зоны скана документа и маркеров шаблона документа;

3. сегментация информационной зоны скана документа на множество машинопечатных текстовых информационных зон документа.

## Список литературы

1. *Архипов О.П., Архипов П.О., Зыкова З.П., Носова Н.Ю.* Применение машиночитаемых зон в современных информационных системах // *Современные технологии безопасности*, 2006. № 3(18)–4(19). С. 32–34.
2. *Архипов О.П., Архипов П.О., Зыкова З.П.* Метод генерации цветных машиночитаемых зон в отпечатках офисных принтеров // *Информационные технологии*, 2005. № 11. С. 37–44.
3. *Архипов О.П., Зыкова З.П., Архипов П.О.* Программа для ЭВМ «Комплекс цветного штрихового кода (СВС)». – Свидетельство № 2005613093 от 28.11.2005 г.
4. *Архипов О.П., Зыкова З.П., Архипов П.О., Зыков Р.В.* Программа для ЭВМ «Автоматизированная система контроля доставки документов строгого учета на бумажном носителе (АС КДДБН)». – Свидетельство № 2007610085 от 09.01.2007 г.
5. *Архипов О.П., Архипов П.О., Зыкова З.П., Захаров В.Н.* Информационная технология структурирования цветовых пространств периферийных устройств при кодировании данных в цветных графических изображениях // *Наукоёмкие технологии*, 2004. Т. 5. №7. С. 43–49.
6. *Архипов О.П., Архипов П.О., Зыкова З.П., Захаров В.Н.* Использование цвета для повышения информационной емкости штрих-кода // *Наукоёмкие технологии*, 2005. Т. 6. № 6. С. 76–82.
7. *Архипов О.П., Зыкова З.П., Архипов П.О.* Программа для ЭВМ «Комплекс офисного цифрового штампа (ODS)». – Свидетельство № 2005612639 от 11.10.2005 г.

# Метод фильтрации сегментированных изображений

А.И. Сорокин<sup>49</sup>

**Аннотация:** В статье представлен один из подходов фильтрации изображений. Предлагаемый метод основан на работе с изображением, разбитым на сегменты, с помощью модифицированной цветовой палитры. Представлен алгоритм фильтрации изображения.

**Ключевые слова:** фильтрация; сегментация; палитра; цветовые координаты; сегменты; Евклидова метрика; пороговое значение; пиксель

## Введение

Создание трехмерных моделей объектов, в связи с развитием мультимедийных ресурсов, является весьма актуальной задачей. Быстрое и качественное получение 3D модели объекта возможно при использовании принципов стереоскопии [1, 2]. Данный подход позволяет извлечь информацию из двумерных изображений объекта для получения трехмерных координат.

Извлечение такой информации возможно при помощи сегментации[3]. Для того чтобы сегментация изображений была качественной необходимо осуществить устранение помех, присутствующих на нем. Помехи бывают различного характера. Это связано с тем, что искажения вносятся еще на этапе съемки, как самим прибором съемки, так и средой, в которой была осуществлена съемка – например неравномерное освещение. Помехи такого рода называются цифровым шумом [5]. Так же, помехи возникают из-за изменения палитры изображения, например, уменьшения количества цветов на изображении, с целью сделать его более контрастным. Такие помехи вносят сложности в обработку изображений. Чтобы уменьшить влияние помех на обработку изображения прибегают к фильтрации.

---

<sup>49</sup> ФИЦ ИУ РАН (Орловский филиал); 1987 г.р.; webdi@mail.ru

При фильтрации изображения, цвет пикселя или группы пикселей заменяется таким, которое не мешает воспринимать и анализировать изображение, то есть, значением, которое не идентифицируется как шум.

Существует много различных методов и алгоритмов фильтрации. Это связано с тем, что используется большое количество критериев оптимальности при обработке изображений, которые выводятся из конкретной задачи.

Все множество подходов к фильтрации изображений делится на две большие категории: пространственные и частотные методы фильтрации [6, 7]. При частотной фильтрации изображений к шумам – сигналам применяются ортогональные преобразования с целью их последующей модификации. Данная категория подходов требует следующей последовательности преобразований:

- преобразование изображения из пространственной области в частотную (дискретное преобразование Фурье),
- осуществление преобразований в частотном спектре,
- обратное преобразование изображение из частотного спектра в пространственную область.

Данный вид методов хорошо справляется с зашумлением изображений, за счет математического аппарата, лежащего в его основе. Недостатки, связанные с этим методом, заключаются в том, что необходимо осуществлять трудоемкий и ресурсоемкий процесс преобразований изображения.

Методы пространственной фильтрации объединяют подходы, в основе которых лежит прямое манипулирование пикселями изображения [6, 7]. Он менее вариативен и более прост, но за счет меньшего объема вычислений обеспечивает достойный результат, затрачивая на это меньше ресурсов.

Целью рассматриваемого в статье исследования является разработка метода фильтрации шумов на изображении, который относится к методам пространственной фильтрации и осуществляется в два этапа.

## 1. Описание метода фильтрации

Первоначально осуществляется подготовка изображения, которая заключается в уменьшении количества цветов на изображении. Далее, после сегментации, алгоритм которой подробно описан в статье [8], осуществляется уменьшение зашумляющих изображение сегментов.

На первом этапе осуществляется уменьшение цветов на изображении. Это, так называемый, процесс «загрубления» цветовых характеристик изображения. За счет уменьшения цветов на изображении, происходит слияние мелких областей в несколько пикселей с крупными областями. Цвета изображения приводятся к цветам уменьшенной палитры. Данный процесс представлен в статье [9]. В процессе выполнения алгоритма, происходит объединение нескольких цветов, лежащих в одном цветовом диапазоне. Размер цветового диапазона определяется евклидовой метрикой, значение которой сравнивается с пороговым значением. Для этого происходит сравнение каждого пикселя изображения с уменьшенной палитрой, которая содержит 4096 цветов. В результате чего увеличиваются цветовые перепады между частями изображения.

Затем, после сегментации, осуществляется пространственная обработка изображения, которая заключается в объединении цветовых сегментов. Данный процесс определяется формулой (1):

$$f(x, y) = P[q(x, y)], \quad (1)$$

где  $q(x, y)$  – изображение, искаженное помехами,

$f(x, y)$  – обработанное изображение,

$P$  – оператор преобразования над  $f$ ,

Если говорить более детально, то пространственная фильтрация изображения размер  $L \times H$  выполняется как операция преобразования двумерной характеристики фильтра  $k(s, t)$  с изображением  $f(x, y)$  [10], представлена формулой(2),

где  $s$ – координата характеристики в горизонтальном направлении вдоль оси  $x$ ,  $s \in [-L / 2, L / 2]$ ,  $t$  – координата характеристики в вертикальном направлении вдоль оси  $y$ ,  $t \in [-H / 2, H / 2]$ :

$$q(x, y) = f(x, y) \times k(x, y) = \sum_{t=-\frac{L}{2}}^{\frac{L}{2}} \sum_{s=-\frac{H}{2}}^{\frac{H}{2}} f(s, t) k(x - s, y - t) = \sum_{t=-\frac{L}{2}}^{\frac{L}{2}} \sum_{s=-\frac{H}{2}}^{\frac{H}{2}} f(x - s, y - t) k(s, t) \quad (2)$$

Так как пиксели изображения сгруппированы в сегменты, то над ними легко осуществить необходимые манипуляции, благодаря которым количество присутствующих шумов уменьшается. В общем виде, описанные действия можно представить формулой (3):

$$f(x, y) = \sum_{i=0}^L \sum_{j=0}^H V(i, j) q(x, y) \quad (3)$$

где  $V(i, j)$  – функция, задающая правила преобразования над изображением  $q(x, y)$ .

Алгоритм объединения цветowych сегментов использует следующие обозначения:

$N$  – переменная целого типа, показывающая количество пикселей, принадлежащего шуму.

$EnvList$  – список, содержащий, цвета и координаты пикселей сегментов, окружающих текущий сегмент. В процессе работы алгоритма, осуществляется просмотр всех сегментов, в которых число пикселей менее  $N$ . Для таких сегментов осуществляется просмотр и подсчет окружающих их - граничных пикселей. В результате, текущему сегменту присваивается цвет пикселей, которых в окружении больше.  $N$  рассчитывается по формуле (4).

$$N = \frac{L \times H}{L + H} \times C \quad (4)$$

где  $L$  – ширина изображения,  $H$  – высота изображения,  $C$  – константа, равная 0,35.

Алгоритмом объединения цветowych сегментов выполняется следующая последовательность действий.

1. Просмотр всех сегментов изображения.
2. Если количество пикселей в текущем сегменте менее  $N$ , то шаг 3 иначе 1.
3. Просмотр граничных пикселей и идентификация окружения текущего сегмента.

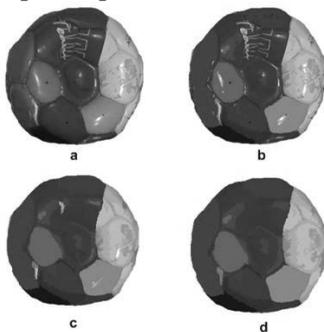
4. Занесение информации об окружающих пикселях в EnvList.

5. Просмотр элементов списка EnvList, с целью осуществления подсчета пикселей каждого цвета.

6. Присвоить текущему сегменту показатель цвета, пикселей, число которых в окружении максимально.

## 2. Результаты работы алгоритма

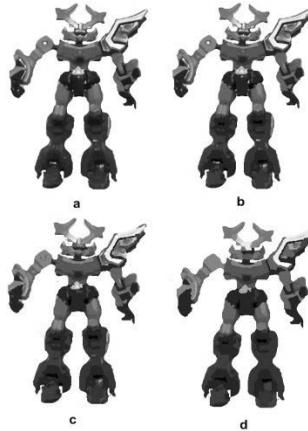
Для демонстрации работы представленного метода фильтрации были выбраны следующие изображения: «Мяч» – Рис. 1, изображение размером 350×350 пикселей; «Робот» – Рис. 2, изображение размером 300×400 пикселей; «Манекен» – Рис. 3, изображение размером 750×1100 пикселей.



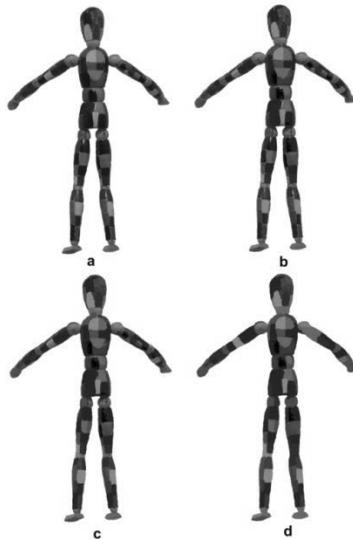
**Рис. 1.** Результаты фильтрации изображения «Мяч»: а – исходное изображение; б – результат работы алгоритма объединения близких цветовых координат; с – при  $N=50$ ; d – при  $N=90$

Первоначальная фильтрация была осуществлена на этапе объединения близких цветовых координат (Рис. 1b). Далее, после сегментации изображения, за счет объединения мелких сегментов, с более крупными, процесс исчезновения шумов представлен более наглядно.

На Рис. 2 представлен результат обработки изображения «Робот», размером 300×400 пикселей.



**Рис. 2.** Результаты фильтрации изображения «Робот»: а – исходное изображение; б –объединенные цветовые координаты; с – при  $N=50$ ; д – при  $N=90$



**Рис. 3.** Результаты фильтрации изображения «Манекен»: а – исходное изображение; б –объединенные цветовые координаты; с – при  $N=100$ ; д – при  $N=200$

На Рис. 3 представлен результат обработки изображения «Манекен», размером  $750 \times 1100$  пикселей.

## Список литературы

1. *Архипов О. П., Маньяков Ю. А., Сиротинин Д. О.* Метод генерации виртуальной сетки опорных точек на цветных изображениях // Информационные технологии в науке, образовании и производстве (ИТНОП): Материалы V Международной научно-технической конференции (Орел, 17–18 мая 2012). – Орел: Госуниверситет-УНПК, 2012.
2. *Архипов О. П., Маньяков Ю. А., Сиротинин Д. О., Сорокин А. И.* Информационная модель метода пофрагментного анализа и представления изменений натурального объекта // Информационные системы и технологии, 2012. № 1. С. 17–24.
3. *Шапиро Л., Стокман Дж.* Компьютерное зрение – Бином, 2009. – 763 с.
4. *Форсайт Д., Понс Ж.* Компьютерное зрение. Современный подход Computer Vision: A Modern Approach. – М.: Вильямс, 2004. – 928 с.
5. [https://ru.wikipedia.org/wiki/Цифровой\\_шум\\_изображения](https://ru.wikipedia.org/wiki/Цифровой_шум_изображения) – *Цифровой шум изображения.*
6. *Гонсалес Р., Вудс Р.* Цифровая обработка изображений. – М.: Техносфера, 2005. – 1072с.
7. *Гонсалес Р., Вудс Р., Эддинс С.* Цифровая обработка изображений в среде MATLAB. Москва: Техносфера, 2006. – 616с.
8. *Сорокин А.И.* Алгоритм сегментации изображений натуральных объектов // Сборник докладов 3-ей Школы молодых ученых ИПИ РАН: 11 декабря 2012 г. - Москва, ИПИ РАН, 2012. С. 55–66.
9. *Архипов П.О., Сорокин А.И.* Информационная технология формирования характеристических признаков предварительной идентификации цветных изображений // Информационные системы и технологии, 2012. №6. С 120–124.
10. *Фисенко В.Т., Фисенко Т.Ю.* Компьютерная обработка и распознавание изображений учебное пособие – ИТМО СПб, 2008.

# Самосинхронный умножитель для распределенных вычислений

Б.А. Степанов<sup>50</sup>, Ю.Г. Дьяченко<sup>51</sup>, Ю.В. Рождественский<sup>52</sup>

**Аннотация:** Доклад посвящен особенностям реализации самосинхронного умножителя для устройства умножения-сложения. Обоснован выбор оптимального самосинхронного кодирования сигналов умножителя с помощью избыточного троичного самосинхронного кода. Предложена структура дерева Уоллеса – основной части умножителя. Рассматриваются проблемы схмотехнической и топологической реализации умножителя. Представлен вариант реализации самосинхронного умножителя в КМОП технологии с проектными нормами 65 нм.

**Ключевые слова:** самосинхронная схема; троичное кодирование; умножитель; дерево Уоллеса

## Введение

Успешное развитие информационных технологий во многом обеспечивается вычислительной мощностью средств ее обработки. Одним из наиболее значимых функциональных блоков современных вычислителей является умножитель со сложением, выполняющий две последовательные операции: умножение двух операндов и сложение с третьим операндом, без округления промежуточных результатов (Fused Multiply-Add, FMA). Это обеспечивает более высокое быстродействие и лучшую точность вычислений.

Самым затратным функциональным блоком устройства FMA, определяющим его потребительские характеристики, является умножитель. Поэтому разработка его оптимальной СС реализации является актуальной задачей.

Доклад посвящен проблемам разработки и реализации

---

<sup>50</sup> ФИЦ ИУ РАН (ИПИ РАН); 1994 г.р.; gtx360@mail.ru

<sup>51</sup> ФИЦ ИУ РАН (ИПИ РАН); 1958 г.р.; diaura@mail.ru

<sup>52</sup> ФИЦ ИУ РАН (ИПИ РАН); 1954 г.р.; YRogdest@ipiran.ru

(схемотехнической и топологической) СС умножителя для устройства, выполняющего операцию.

## 1. СС умножитель

В литературе известны многочисленные реализации синхронного умножителя. Однако наилучшее сочетание потребительских характеристик умножителя обеспечивается применением СС схем, не зависящих от задержек элементов (Speed Independent, SI), которые обладают более широким диапазоном работоспособности и, в отличие от синхронных схем, свойством самодиагностики. В своих исследованиях [1] авторы подтвердили, что грамотно спроектированные SI схемы обладают также лучшим быстродействием и энергопотреблением в сравнении с синхронными аналогами.

В соответствии со стандартом IEEE 754 на умножитель поступают два 64-разрядных операнда. Фактически умножаются только мантиссы этих операндов, экспоненты складываются отдельно. Наилучшее решение, с точки зрения аппаратурных затрат и быстродействия, обеспечивается использованием модифицированного алгоритма Бута.

В результате умножитель состоит из кодера Бута и сумматора формируемых им частичных произведений. Схемотехническая реализация умножителя и его характеристики во многом зависят от типа кодирования входных операндов и промежуточных результатов

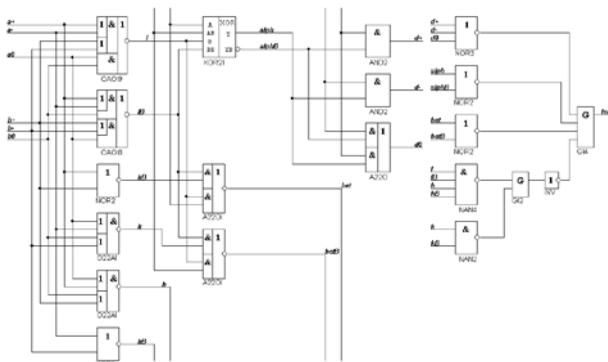
### 1.1 СС кодирование сигналов в умножителе

В качестве прототипа алгоритмического решения дерева Уоллеса была выбрана синхронная реализация [2], использующая избыточное кодирование операндов. После проведенного анализа возможных СС кодов авторами была предложена схема кодирования, представленная в работе [3, Таблица 1]. Такая избыточность позволяет описывать сумму двух однобитных чисел в прямом и обратном кодах одним кодовым чис-

лом и производить операцию сложения в одном кодовом разряде с сохранением переноса. В результате сокращается количество промежуточных и выходных сигналов и уменьшается число каскадов дерева Уоллеса.

## 1.2 Схмотехника СС умножителя

Основные аппаратурные затраты СС умножителя определяются сумматором частичных произведений, реализуемого деревом Уоллеса, которое строится из одноразрядных сумматоров. На рис. 1 показана схема одноразрядного СС сумматора с троичными выводами. В правой части показана подсхема сто-процентной индикации сумматора.



**Рис. 1.** Троичный СС сумматор

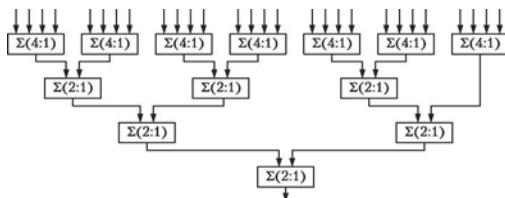
Сравнение показывает, что СС сумматор с парафазными выводами [4] имеет меньшую сложность в сравнении с троичным СС сумматором, но при построении умножителя на его основе получается схема и более сложная, и менее быстродействующая, чем схема умножителя на троичном сумматоре.

## 1.3 Дерево Уоллеса СС умножителя

Реализация дерева Уоллеса для сложения операндов, кодированных алгоритмом Бута, для 53-разрядных чисел,

представленных в троичном СС коде, показана на рис. 2. Его входами являются 27 частичных произведений, формируемых кодером Бута, и некоторое корректирующее слагаемое.

За счет СС троичного кодирования внутренних сигналов, предложенной в [2] методики преобразования каждого второго частичного произведения и учета дополнительного корректирующего частичного произведения первый каскад дерева Уоллеса обеспечивает сжатие 27 частичных произведений до 7 троичных сигналов. На следующих каскадах дерева Уоллеса троичные сумматоры обеспечивают коэффициент сжатия числа сигналов 2:1, реализуя наиболее быстрое сложение практически без аппаратной избыточности.



**Рис. 2.** СС реализация дерева Уоллеса

Использование троичных СС сумматоров обеспечивает повышение быстродействия дерева Уоллеса по сравнению с классическим СС алгоритмом более чем на 20%. Снижение аппаратных затрат также составило около 20%.

## 2. Топологическая реализация

Одна из основных проблем СС схем – большое количество сигналов, т.к. СС коды используют избыточное представление информации. Для парафазного кодирования получается удвоение числа сигналов. Поэтому при топологическом проектировании многоразрядных СС устройств, как правило, целесообразно использовать структуры с вертикальным распространением сигналов от блока к блоку.

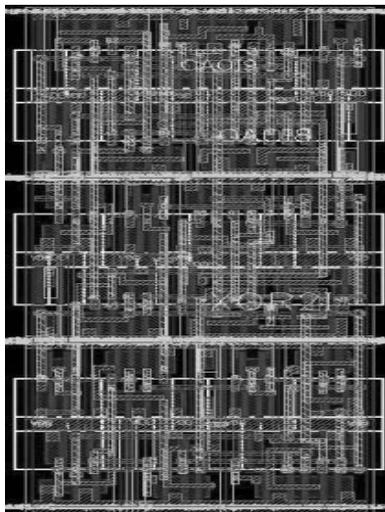
Троичное кодирование операндов в СС умножителе

смягчает данную проблему: дерево Уоллеса является совокупностью сумматоров с сохранением переноса и при парафазном кодировании операндов каждый сумматор в умножителе имеет 4 выхода, а при троичном – только 3 выхода.

Контур топологии умножителя целесообразно делать близким к прямоугольнику с отношением сторон 0,3 ... 1. Исходя из этого, были выбраны габариты топологии одноразрядного сумматора.

## 2.1 Топология одноразрядного сумматора

Рис. 3 демонстрирует топологию одноразрядного троичного СС сумматора. Здесь троичные входы располагаются сверху, выходы суммы – снизу, а сигналы межразрядного переноса – сбоку. Это облегчает стыковку соседних одноразрядных сумматоров и обеспечивает распространение информации в вертикальном направлении.



**Рис. 3.** Топология троичного сумматора

Топология одноразрядного сумматора, реализованного по стандартной КМОП технологии с 65-нм проектными нормами и 9 слоями металлизации, имеет размеры  $4,56 \times 12,0$  мкм.

Для разводки использованы только второй и третий слои металлизации, что обеспечивает его «прозрачность» для трассировки глобальных сигналов.

## 2.2 Топология умножителя

Весь умножитель, топология которого изображена на рис. 4, имеет размеры  $570 \times 234$  мкм. Здесь также входные операнды поступают сверху, выходы расположены снизу. Индикаторные сигналы на выходе блока умножителя обеспечивают его поразрядную индикацию.

Умножитель был спроектирован как часть устройства умножения-сложения. В настоящее время ведется работа по встраиванию его в топологию всего блока умножения-сложения.

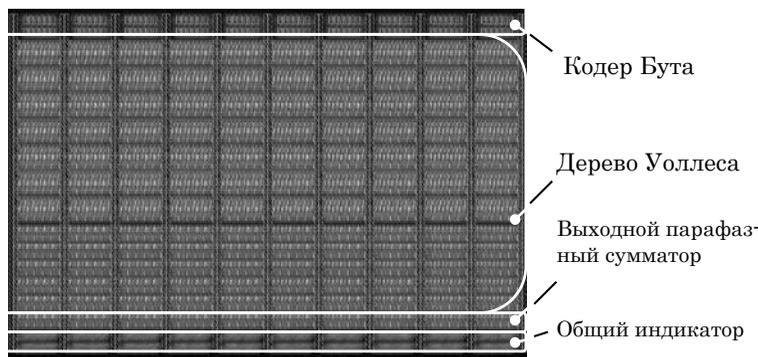


Рис. 4. Топология умножителя

## Выводы

Самосинхронные схемы обладают рядом весомых преимуществ по сравнению с синхронными аналогами: низкое энергопотребление, самодиагностика, высокая скорость работы.

Однако они отличаются сложностью разработки. К тому же, самосинхронное кодирование увеличивает количество сигналов в схеме, что приводит к увеличению аппаратных затрат.

Переход к троичному кодированию операндов повышает эффективность использования алгоритма дерева Уоллеса, повышая степень сжатия операндов в нем, уменьшает аппаратные затраты на реализацию вычислителя в сравнении с парафазным кодированием.

Структура топологии с вертикальным распространением сигналов позволяет упростить трассировку сигналов в схеме. При этом в топологии входы и выходы переносов одноразрядного сумматора распространяются горизонтально.

## Список литературы

1. *Степченко Ю. А., Дьяченко Ю. Г., Горелкин Г. А.* Самосинхронные схемы – будущее микроэлектроники. // Вопросы радиоэлектроники, 2011. No.2. С. 153–184.
2. *Makino H., Nakase Y., Suzuki H., Morinaka H., Shinohara H., and Mashiko K.* An 8.8 ns 54x54 bit Multiplier with High Speed Redundant Binary Architecture // IEEE Journal of Solid-State Circuits, 1996. V.31 No.6. P. 773–783.
3. *Соколов И. А., Степченко Ю. А., Бобков С. Г., Рождественский Ю. В., Дьяченко Ю. Г.* Умножитель с накоплением: методологические аспекты // Системы и средства информатики, 2014. Т. 24. No.3. С. 44–62.
4. *Плеханов Л. П.* Основы самосинхронных электронных схем. – М.: Бином. Лаборатория знаний, 2013. – 208 с.

# Модель для анализа вероятности прерывания обслуживания в сети 3GPP LTE с системой LSA\*

А.А. Федуро<sup>53</sup>, В.Р. Филипова<sup>54</sup>, Е.В. Маркова<sup>55</sup>,  
Е.В. Мокров<sup>56</sup>, И.А. Гудкова<sup>57</sup>

**Аннотация:** Современное развитие информационных технологий тесно связано с проблемой стремительного роста объема передаваемых данных в условиях ограниченности частотного диапазона мобильных сетей. Одним из возможных решений данной проблемы является использование разрабатываемой при поддержке Европейской комиссии (European Commission, EC) системы совместного использования лицензированного спектра частот (Licensed Shared Access, LSA). В работе построена модель схемы совместного использования радиочастот оператором, имеющим доступ как к полосе совместного использования, так и индивидуальной полосе, проведен анализ основных вероятностных характеристик модели – вероятности блокировки и вероятности прерывания.

**Ключевые слова:** мобильная сеть; совместное использование радиочастот; LSA; инфинитезимальная матрица; вероятность блокировки; вероятность прерывания

## Введение

Согласно данным, опубликованным компанией Cisco Systems [1], к 2019 году объем трафика, передаваемого в мобильных сетях, будет составлять около 24.3 эксабайт в месяц. В связи с этим для обслуживания пользователей с требуемым

---

\*Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов № 13-07-00953 и 15-07-03608.

<sup>53</sup> РУДН; 1994 г.р.; feduro.nastya@yandex.ru

<sup>54</sup> РУДН; 1993 г.р.; valeryflp@gmail.com

<sup>55</sup> РУДН; 1987 г.р.; к.ф.-м.н.; mkatyushka@mail.ru

<sup>56</sup> РУДН; 1988 г.р.; melkor77@yandex.ru

<sup>57</sup> РУДН; 1985 г.р.; к.ф.-м.н.; igudkova@sci.pfu.edu.ru

уровнем QoS (Quality of Service, QoS) у мобильных операторов возникает необходимость поиска дополнительных ресурсов. В настоящее время для решения этой проблемы при поддержке Европейской комиссии разрабатывается система совместного использования частотного спектра LSA.

В международных стандартах правила совместного использования операторами радиочастот не определены, поэтому возможны различные варианты моделей схем сетей четвертого, пятого поколений [2], в частности сетей на базе технологии LTE [3] с применением системы LSA [4]. Схемы могут отличаться друг от друга, в первую очередь, числом операторов, типом приоритета их доступа к полосе, например, оператор с большим приоритетом – владелец полосы совместного использования, оператор с меньшим приоритетом – арендатор [4], алгоритмами распределения нагрузки между индивидуальной полосой и полосой совместного использования. Важным отличием схем являются методы управления радиоресурсами [5] – политики реагирования оператора-арендатора на необходимость доступа к полосе совместного использования оператора-владельца.

В работе построена модель схемы совместного использования радиочастот оператором-арендатором, имеющим доступ, как к полосе индивидуального использования емкостью  $C_1$ , так и совместной полосе емкостью  $C_2$ . Управление радиоресурсами основано на реализации «жесткой» политики, т.е. при необходимости полоса совместного использования радиочастот переходит оператору с более высоким приоритетом – владельцу. Отсутствие явных аналитических формул, необходимых для расчета показателей качества обслуживания модели, приводит к необходимости численного решения системы уравнений равновесия (СУР).

## 1. Построение модели

Предположим, что входящий поток является пуассоновским с параметром  $\lambda$ , время обслуживания распределено по

экспоненциальному закону со средним  $\mu^{-1}$ . Обозначим  $\alpha$  – интенсивность отключения полосы совместного использования, а  $\beta$  – интенсивность восстановления полосы,  $C = C_1 + C_2$  – общую емкость системы.

Пусть  $n_1$  – количество заявок на полосе индивидуального использования, а  $n_2$  – на полосе совместного использования,  $s$  – состояние полосы совместного использования ( $s = 0$ , если полоса выключена, и  $s = 1$ , если полоса включена). Состояние системы в некоторый момент времени описывает вектор  $(n_1, n_2, s)$  в пространстве состояний

$$X = \left\{ \begin{array}{l} n_1 = 0, \dots, C_1, n_2 = 0, \dots, C_2, s = 1 \\ n_1 = 0, \dots, C_1, n_2 = 0, s = 0 \end{array} \right\} \quad (1)$$

При поступлении заявки на обслуживание возможны три случая:

1. заявка обслуживается на полосе индивидуального использования, если количество обслуживаемых на этой полосе заявок меньше, чем емкость полосы, т.е.  $n_1 < C_1$ ;
2. заявка обслуживается на полосе совместного использования, если полоса индивидуального использования полностью занята, и количество обслуживаемых заявок на полосе совместного использования меньше, чем емкость этой полосы, т.е.  $n_1 = C_1$  и  $n_2 < C_2$ ;
3. заявка блокируется, если обе полосы заняты, т.е.  $n_1 + n_2 = C$ .

При отключении полосы совместного использования заявки, обслуживаемые на ней, переходят на полосу индивидуального использования, если на ней есть свободные ресурсы.

## 2. Инфинитезимальная матрица

Стационарное распределение  $P(n_1, n_2, s)$  вероятностей





точки зрения пользователя  $I_{cus}$  – вероятность того, что обслуживание конкретного пользователя будет прервано, – и с точки зрения оператора  $I_{op}$  – вероятность того, что обслуживание хотя бы одного пользователя будет прервано. Эти характеристики рассчитываются по формулам (2), (3), (4) соответственно:

$$B = p(C_1, 0, 0) + p(C_1, C_2, 1), \quad (2)$$

$$I_{cus} = \sum_{n_2=1}^{C_2} \sum_{n_1=C_1-n_2+1}^{C_1} p(n_1, n_2, 1) \cdot \frac{\alpha}{\alpha + (n_1 + n_2)\mu + \lambda \cdot 1_{\{C_1-n_1+s(C_2+n_2)>0\}}}, \quad (3)$$

$$I_{op} = \sum_{n_2=1}^{C_2} \sum_{n_1=C_1-n_2+1}^{C_1-1} \frac{n_2 - C_1 + n_1}{n_2} p(n_1, n_2, 1) \cdot \frac{\alpha}{\alpha + (n_1 + n_2)\mu + \lambda} + \sum_{n_2=1}^{C_2-1} p(C_1, n_2, 1) \cdot \frac{\alpha}{\alpha + (n_1 + n_2)\mu + \lambda} + p(C_1, C_2, 1) \cdot \frac{\alpha}{\alpha + (n_1 + n_2)\mu}, \quad (4)$$

Для проведения численного анализа рассмотрим систему, в которой общая ширина полос составляет 2,5 МГц (см. Рис. 1). Передача данных осуществляется на гарантированной скорости, равной 250Кбит/с. С учетом спектральной эффективности LTE, на каждой из двух полос может быть обслужено до пяти пользователей. Интенсивности выключения и восстановления полосы совместного использования соответственно равны  $\alpha = 1 / 23$  и  $\beta = 1 / 37$ . Рассматривается случай передачи данных беспроводной камерой, не имеющей собственного буфера фотоснимка высокого качества.

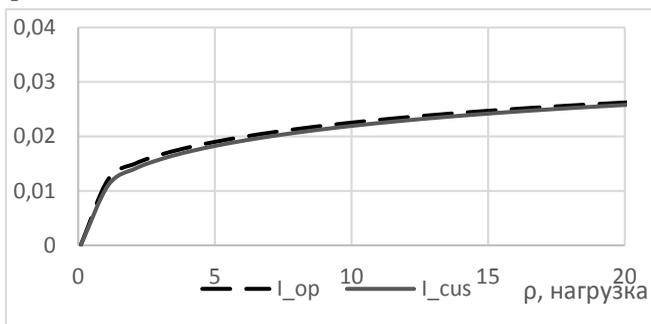


Рис. 1. Вероятности прерывания  $I_{cus}$  и  $I_{op}$

## Выводы

В работе построена модель схемы совместного использования радиочастот, получен блочный вид инфинитезимальной матрицы и представлено численное решение СУР. Проведен численный анализ вероятностей прерывания – с точки зрения пользователя и с точки зрения оператора при различных значениях интенсивности предложенной нагрузки.

## Список литературы

1. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update 2014–2019, February 3, 2015.
2. *Shorgin S.Y., Samouylov K.E., Gudkova I.A., Galinina O.S., and Andreev S.D.* On the benefits of 5G wireless technology for future mobile cloud computing // Proc. of the 1st International Science and Technology Conference «Modern Networking Technologies (MoNeTec): SDN & NFV – The Next Generation of Computational Infrastructure». – Moscow, MAKS Press., 2014. P. 151–154.
3. *Stasiak M., Glabowski M., Wisniewski A., and Zwierzykowski P.* Modelling and dimensioning of mobile wireless networks: from GSM to LTE. – Willey, 2010. – 340 p.
4. *Borodakiy V.Y., Samouylov K.E., Gudkova I.A., Ostriкова D.Y., Ponomarenko A.A., Turlikov A.M., and Andreev S.D.* Modeling unreliable LSA operation in 3GPP LTE cellular networks // Proc. of the 6th International Congress on Ultra Modern Telecommunications and Control Systems ICUMT-2014. – IEEE., 2014. P. 490–496.
5. *Borodakiy V. Y., Gudkova I. A., Markova E. V., and Samouylov K. E.* Modelling and performance analysis of pre-emption based radio admission control scheme for video conferencing over LTE // Proc. of the 2014 ITU Kaleidoscope Academic Conference. – 2014. P. 53–59.

# Исследование возможностей прогнозирования нефтедобычи с помощью математического моделирования\*

А. Ю. Флёрова<sup>58</sup>

**Аннотация:** В данной работе построена математическая модель процесса управления разработкой нефтяного месторождения, с помощью которой исследованы возможности прогноза количества извлекаемой.

**Ключевые слова:** нефтедобыча; задача оптимального управления

## Введение

В середине прошлого века М.К. Хубберт предложил подход к количественному прогнозированию добычи нефти в США и во всем мире, позволивший предсказать пик нефтедобычи в США в начале 70-х годов [1]. Этот подход был основан на том, что нефть является конечным ресурсом, динамика извлечения и разведки которого может быть описана в форме определенных кривых. Производная логистической кривой, называемая кривой Хубберта, имеющая колоколообразную форму, наиболее близко, по мнению Хубберта, отражала реальный ход добычи нефти. Уравнение этой кривой имеет вид

$$\dot{X}(t) = V \frac{X(t)}{Y} (Y - X(t)),$$

где  $X(t)$  – накопленная добыча в момент времени  $t$ ,  $Y$  – накопленные извлекаемые запасы,  $V$  – постоянная величина, характеризующая темп роста накопленной добычи. Однако, к концу

---

\*Работа выполнена при поддержке РФФИ (проекты №№ 16-57-53078, 14-07-00075).

<sup>58</sup> ФИЦ ИУ РАН (ВЦ им. А. А. Дородницына РАН); 1981; к.ф.-м.н.; a.flerova@mail.ru

прошлого века кривые Хубберта утратили свою точность в прогнозе добычи, и в данной работе сделана попытка модифицировать кривые Хубберта для получения более точных прогнозов с учетом бурного развития технологий.

## 1. Математическая модель

Введем функцию  $b(t)$ , описывающую технологический уровень процесса нефтедобычи, приращением которой в единицу времени мы можем управлять с помощью неотрицательного управления  $u(t)$ . Время разработки месторождения  $T$  не фиксировано и является в задаче определяемым параметром. Через  $x(t)$  обозначим долю извлекаемых запасов, примем объем месторождения за единицу,  $0 \leq x(t) \leq 1$ , уровень начала добычи  $x(0) = x_0$  не фиксирован. Пусть производитель максимизирует свой доход за время жизни месторождения. В данной модели цены на нефть считаем постоянными. Для того чтобы учесть возможные инфраструктурные ограничения на добычу, введем параметр  $M$ . Получаем следующую задачу оптимального управления с фазовыми ограничениями

$$\begin{aligned} N(T) - cb(0) &\rightarrow \max \\ \dot{N}(t) &= (v(t) - cu(t) - c_0)e^{-rt}, \\ \dot{x}(t) &= v(t), \\ \dot{b}(t) &= u(t), \\ v(t) &\leq b(t)x(t)(1 - x(t)), \\ 0 &\leq v(t) \leq M, u(t) \geq 0, \\ N(0) &= 0. \end{aligned}$$

Здесь  $c$  и  $c_0$  – параметры, определяющие стоимость добычи нефти (издержки, связанные с введением новых технологических мощностей, и постоянные издержки, связанные с эксплуатацией месторождения, соответственно),  $r$  – коэффициент дисконтирования,  $a b(0)$  – начальный технологический уровень (необходимые затраты перед началом добычи).

В этой задаче управляющими параметрами являются  $u(t)$  и  $v(t)$ . Первый характеризует рост технологических возможностей, а второй – текущую добычу нефти. Для такой задачи сформулирован и доказан принцип максимума Понтрягина [2], применение которого приводит описанию различных режимов управления, зависящих от параметров модели.

## 2. Оптимальное управление

Применение принципа максимума дает нам три возможных варианта поведения оптимальной траектории. Далее функции  $\varphi_2(t)$  и  $\varphi_3(t)$  – сопряженные переменные, для которых выполняются условия транверсальности  $\varphi_2(0) = \varphi_2(T) = \varphi_3(T) = 0$  и  $\varphi_3(0) = c$ .

1. *«Режим Хубберта»*. В этом режиме график функции, характеризующей изменение накопленной добычи, имеет колоколообразную форму, характерную для «традиционных» кривых Хубберта. В этом режиме возможно как начало, так и конец процесса нефтедобычи. В случае хорошо развитой инфраструктуры, т.е. при достаточно большом значении параметра  $M$ , весь процесс нефтедобычи будет проходить по этому сценарию. При этом поведение оптимальных траекторий описывается следующими уравнениями.

$$\begin{aligned}\dot{x}(t) &= v(t) = b(t)x(t)(1 - x(t)) \leq M, \\ \dot{\varphi}_2(t) &= r\varphi_2(t) - (1 + \varphi_2(t))b(t)(1 - 2x(t)), \\ \dot{\varphi}_3(t) &= r\varphi_3(t) - (1 + \varphi_2(t))x(t)(1 - x(t)), \\ b(t) &= \text{const}.\end{aligned}$$

2. *«Режим полочки»*. В этом режиме инфраструктура полностью загружена и добыча происходит на уровне  $M$ . В таком режиме невозможно начинать и заканчивать процесс нефтедобычи, и будут справедливы следующие уравнения.

$$\dot{x}(t) = v(t) = M < b(t)x(t)(1 - x(t)),$$

$$\dot{\varphi}_2(t) = r\varphi_2(t), \dot{\varphi}_3(t) = r\varphi_3(t),$$

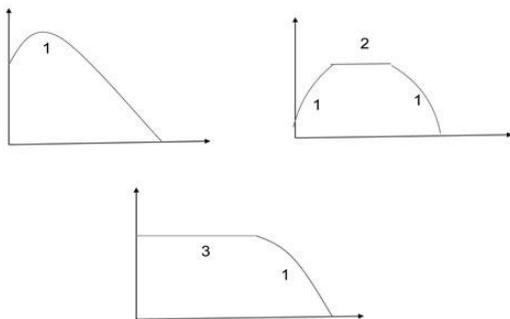
$$b(t) = \text{const.}$$

3. *Режим «особенный».* В этом режиме начальные капиталовложения  $b(0)$  выбираются таким образом, что  $b(0)x(0)(1 - x(0)) = M$ , а управление  $u(t)$  такое, что  $b(t)x(t)(1 - x(t)) = M$ , т.е.

$$u(t) = -\frac{M^2(1 - 2x(t))}{(1 - x(t))^2}.$$

При этом  $x(t) \geq 1/2$ . Выход из этого режима осуществляется через первый режим. Параметры описываются следующими уравнениями.

$$\begin{aligned} \dot{x}(t) &= v(t) = b(t)x(t)(1 - x(t)) = M, \\ \dot{\varphi}_2(t) &= r\varphi_2(t) - \frac{c}{x(t)(1 - x(t))}b(t)(1 - 2x(t)), \\ \varphi_3(t) &= c. \end{aligned}$$



**Рис. 1.** Режимы добычи

Выбор оптимального режима и переходы между режимами определяются в зависимости от начальных параметров модели. Системы дифференциальных уравнений просчитываются численно в обратном времени.

Все возможные режимы представлены на Рис. 1, где по оси абсцисс отложено время, а по оси ординат – изменение накопленной добычи, т.е. текущая добыча. Цифрами отмечены соответствующие возможные режимы.

## **Выводы**

Таким образом, зная исходные данные о месторождении (издержки по добычи нефти, издержки по увеличению технологического уровня, инфраструктурные ограничения) мы можем построить прогноз добычи нефти, найти время жизни месторождения (т.е. время эксплуатации), найти оптимальное количество добываемой нефти (начальный уровень добычи и уровень отсечения), указать необходимый начальный технологический уровень и оптимальное управление развитием технологий.

## **Благодарности**

*Выражаю благодарность профессору Александру Алексеевичу Шананину за руководство данной работой и Виталию Шайдунову за первые идеи.*

## **Список литературы**

1. *Петров В.В., Поляков Г.А., Полякова Т.В., Сергеев В.М.* Долгосрочные перспективы российской нефти (анализы, тренды, сценарии). – М.: ФАЗИС, 2003. – 200 с.
2. *Милютин А.А., Дмитрук А.В., Осмоловский Н.П.* Принцип максимума в оптимальном управлении. – М., Издательство Центра прикладных исследований при механико-математическом факультете МГУ, 2004. – 284 с.

# Особенности программной реализации имитационной модели потоковой рекуррентной архитектуры\*

Д. В. Хилько<sup>59</sup>, Ю. И. Шикунов<sup>60</sup>, Ю. А. Степченков<sup>61</sup>

**Аннотация:** Статья посвящена программной реализации имитационной модели, входящей в состав аппаратно-программных средств моделирования и отладки вычислительного устройства, построенного на основе многоядерной потоковой рекуррентной архитектуры. Приводится описание архитектуры имитационной модели, особенностей программной реализации ключевых компонентов. Рассматривается механизм организации взаимодействия между блоками и компонентами модели, функционирующими параллельно. Также рассматриваются технические возможности текущей версии программой реализации, перспективы их развития. Демонстрируются некоторые этапы функционирования модели на примере вычисления тестового алгоритма.

**Ключевые слова:** потоковая архитектура; имитационное моделирование; рекуррентность

## Введение

В Институте проблем информатики ФИЦ ИУ РАН на протяжении ряда лет ведутся исследования в области нетрадиционных архитектур вычислительных систем и разработки принципиально новых концепций организации вычислительного процесса. Результатом данных исследований является многоядерная потоковая рекуррентная архитектура (МПРА),

---

\*Работа выполнена при частичной финансовой поддержке РФФИ в рамках научного проекта № 13-07-12068

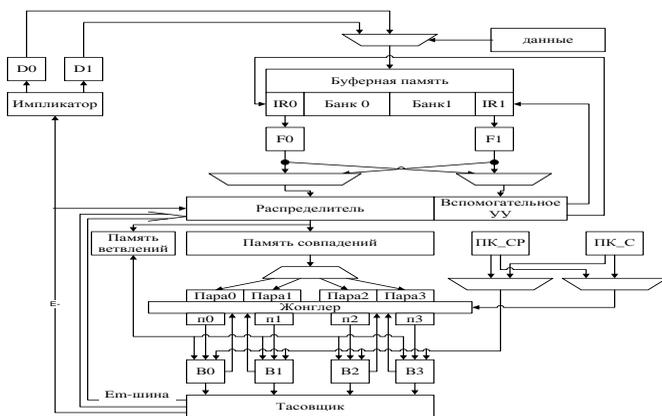
<sup>59</sup> ФИЦ ИУ РАН (ИПИ РАН); 1987 г.р.; dhilko@yandex.ru

<sup>60</sup> ФИЦ ИУ РАН (ИПИ РАН); 1995 г.р.; yishikunov@yandex.ru

<sup>61</sup> ФИЦ ИУ РАН (ИПИ РАН); 1951 г.р.; к.т.н.; ystepchenkov@ipiran.ru

опытный образец которой реализуется в виде гибридной двух-уровневой архитектуры рекуррентного обработчика сигналов (ГАРОС) [1, 2].

В целях обеспечения разработки опытного образца была создана платформа аппаратно-программного моделирования и отладки, названная GAROS IDE [3]. Одним из основных компонентов данной платформы является программная имитационная модель (ИМ) рекуррентного обработчика сигналов (РОС), реализованная в программе СИМПА [4]. На Рис. 1 приведена архитектура имитационной модели РОС.



**Рис. 1.** Архитектура имитационной модели

Указанная модель используется для отладки специализированного ПО, функционирующего в среде ГАРОС, а также аппаратной VHDL-модели. Подробное описание ее составных блоков может быть найдено в статье [5]. Настоящая работа посвящена описанию основных особенностей и возможностей программной реализации разработанной модели.

## 1. Программная реализация модели

### 1.1. Общее описание компонент модели

Существуют несколько основных подходов разработки

ИМ. Для реализации данной ИМ была выбрана дискретно-событийная парадигма моделирования. Суть данного подхода заключается в представлении процесса функционирования системы, как хронологической последовательности событий.

В качестве программной среды реализации ИМ была выбрана платформа *.NET Framework*, а в качестве языка программирования – *C#*. Для каждого компонента модели созданы соответствующие классы, реализующие основные элементы дискретно-событийного подхода (часы, списки действий и событий, набор состояний и т.д.), а также механизм межкомпонентного взаимодействия, описываемый в разделе 1.3.

Разработчики РОС разбили вычислительный процесс на три ступени конвейера. На Рис. 1 можно выделить три основных группы компонентов, образующих указанные ступени конвейера: Буферная память + Распределитель + Вспомогательное УУ (ступень 1), Память совпадений + Память ветвлений + Памяти констант (ступень 2), Вычислители + Тасовщик + Импликатор (ступень 3). Совокупное состояние компонентов ступени определяет состояние этой ступени.

Наиболее функционально сложным компонентом модели является «Распределитель», в задачи которого входит управление входным и промежуточными потоками данных и их преобразование для последующей обработки на ступенях 2 и 3 конвейера. Для данного компонента было разработано большое количество режимов работы, опций конфигурации и пр. Высокая функциональная нагрузка на «Распределитель» является его главной отличительной особенностью от остальных компонент, поэтому в разделе 1.2 приводится краткое описание его программной реализации.

## 1.2. Компонент «Распределитель»

Ввиду особенностей реализации событийной модели в языке *C#*, события модели были названы действиями. Представление состояний на языке *C#*

```
public enum DistributorStates
{
    Stopped = 0, // Начальное состояние S0
    InitState = 1, // Состояние начальной инициализации S1
}
```

```

DataWaiting = 2, // Состояние ожидания данных S2
Grape1Proc = 3, // Состояние обработки первой полуторги S3
Grape2Proc = 4, // Состояние формирования второй полуторги S4
KeyProc = 5, // Состояние формирования горсти ключом S5
MGRapeFinished = 6, // Финальное состояние «Распределителя» S6
Finished = 7, // Финальное состояние обработки капсулы S7
Halted = 8, // Состояние приостановки работы S8
Error = 9 // Состояние останова по ошибке S9
} // Дискретные состояния распределителя

```

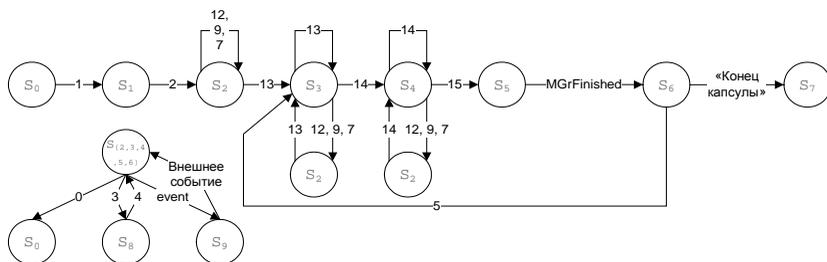
### Представление действий на языке C#

```

public enum DistrActionName
{
    Stop = 0, // Остановить моделирование
    Init = 1, // Инициализировать компонент
    Start = 2, // Старт моделирования
    Hold = 3, // Приостановить моделирование
    Resume = 4, // Возобновить моделирование
    Step = 5, // Выполнить логический шаг
    Configure = 6, // Установить параметры конфигурации
    CallBS = 7, // Запросить данные БП
    WriteFReg = 8, // Записать в регистр F0 или F1
    ReadFReg = 9, // Читать регистр F0 или F1
    UnPack = 10, // Распаковать упакованный операнд
    WriteFIFO = 11, // Записать в FIFO
    ReadFIFO = 12, // Считать из FIFO
    WriteToBranchMemory = 13, // Записать в память ветвлений
    AddToGrape1 = 14, // Добавить к первой полуторги
    AddToGrape2 = 15, // Добавить ко второй полуторги
    MergeEIGrapes = 16 // Выбор или объединение E, I-горстей
} // Внутренние действия распределителя

```

Кроме того, была построена диаграмма состояний компонента «Распределитель», представленная на Рис. 2.



**Рис. 2.** Диаграмма состояний Распределителя

На диаграмме числами указаны номера действий (событий), в результате которых осуществляется переход из одного состояния в другое, а «капсула» – это наименование программы,

исполняемой в модели.

### 1.3. Межкомпонентное взаимодействие

Конвейерная реализации РОС потребовала синхронизировать все его ступени на каждом вычислительном шаге. Для решения этой задачи был применен механизм блокировок. Для этого было разработано три функциональных примитива: *DoReadBusData*, *DoStep*, *DoFinalizeStep* и группа состояний для работы с ними. Данные примитивы реализуют основные точки блокировки, позволяя управлять вычислительными потоками и синхронизировать их на каждом шаге.

В качестве точки старта вычислительного шага выступает состояние *DataReady* (см. Рис. 3), сигнализирующее о готовности данных для обработки. Затем, оно порождает исполнение события *DoReadBusData*, реализующего чтение данных из входных шин компонент, после чего для каждого уровня конвейера выполняется событие *DoStep*. По окончании исполнения все процессы ожидают перехода системы в состояние *Finalize*, после чего исполняется событие *DoFinalizeStep*, реализующее очистку необходимых регистров и запись данных на шины для дальнейшего считывания на следующем шаге.

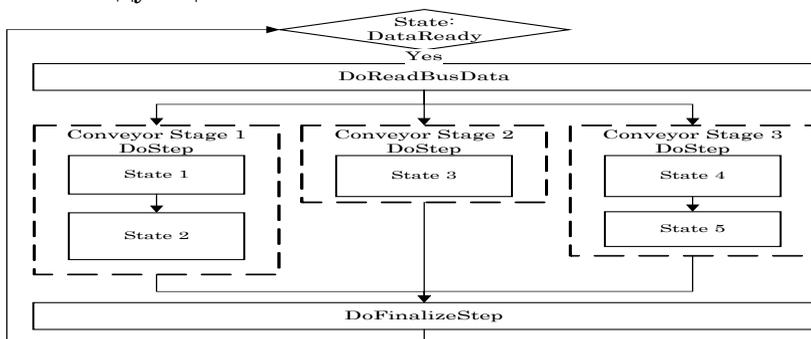


Рис. 3. Механизм межкомпонентного взаимодействия

Данная схема описывает верхний уровень взаимодей-

ствия между ступенями конвейера. В тоже время, если некоторые блоки компонента также могут функционировать параллельно, применяется аналогичная схема взаимодействия.

При разработке ИМ была применена последовательная реализация исполнения параллельных процессов для упрощения отладки логики в каждом отдельном процессе.

## **2. Возможности и перспективы развития ИМ**

Разработанная ИМ входит в состав программных средств платформы GAROS IDE и является одним из ключевых ее компонентов. Текущая версия ИМ обладает следующими техническими возможностями.

1) Позволяет выполнять как одиночный шаг моделирования, так и заданное количество шагов.

2) Имеется возможность осуществлять переход к заданному шагу моделирования (одному из выполненных ранее).

3) Автоматически завершает свое выполнение по достижении условий окончания моделирования, устанавливаемых исполняемой программой.

4) Для каждого сеанса моделирования можно определить степень детализации наблюдаемых результатов.

5) Доступен просмотр состояния любого компонента на каждом шаге моделирования.

6) Визуализация результатов моделирования осуществляется в текстовой форме.

7) Сопоставлять результаты программного моделирования и аппаратного моделирования средствами GAROS IDE.

Достичь подобной гибкости удалось при помощи развитого инструментария ведения лога, хранящего в формализованном виде исчерпывающую информацию о сеансе моделирования для всех компонентов и шагов. В текущей версии логирование осуществляется в соответствии с формируемым пользователем XML-шаблоном, на основе которого строится словарь логируемых сущностей. Сохраняемая информация имеет высокую степень структурированности и формализации.

Потребности разработчиков РОС постоянно возрастают по мере развития и отладки архитектуры. Это обуславливает необходимость расширения функциональных возможностей ИМ и среды GAROS IDE. Имеются следующие перспективы развития.

1) Разработка подсистемы визуализации результатов моделирования, при условии сохранения текстового вывода при необходимости более высокой детализации.

2) Доработка режимов логирования с целью повышения вариативности наблюдений результатов моделирования.

3) Разработка языковых нотаций и подсистем описания новых компонент ИМ с целью превращения ее в полноценную подсистему имитационного моделирования.

### 3. Демонстрация процесса моделирования

Процесс моделирования демонстрирует один из алгоритмов распознавания изолированных слов – полосовая фильтрация, – который был реализован в виде специальной программы: капсулы и отлажен при помощи ИМ. На Рис. 4 приведен завершающий шаг моделирования.

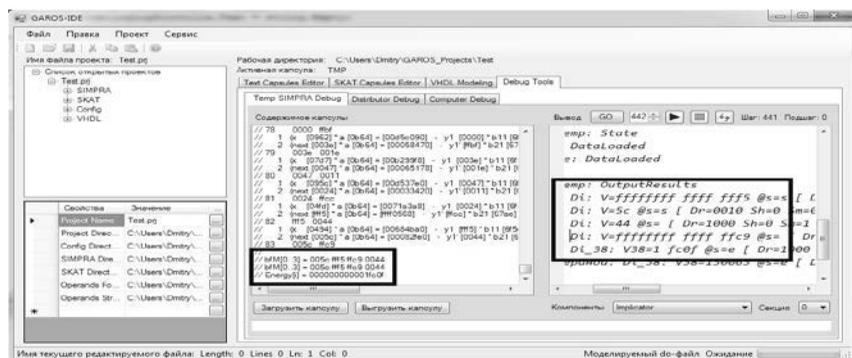


Рис. 4. Завершающий шаг моделирования

## Выводы

В статье рассмотрены особенности программной реализа-

ции ключевых компонент и механизмов имитационной модели РОС. Описаны технические возможности модели, функционирующей в рамках аппаратно-программной платформы GAROS IDE. Также продемонстрирована работа модели на тестовом алгоритме.

В соответствии с потребностями разработчиков определены перспективы развития модели до полноценной подсистемы имитационного моделирования, которая станет мощным инструментарием программного моделирования средствами платформы GAROS IDE.

## Список литературы

1. *Степченков Ю. А., Петрухин В. С.* Особенности гибридного варианта реализации на ПЛИС рекуррентного обработчика сигналов // Системы и средства информатики, 2008. Доп. вып. С. 118–129.
2. *Хилько Д. В., Степченков Ю. А.* Модель потоковой архитектуры на примере распознавателя слов // Системы и средства информатики. 2012, Т. 22. Вып. 2. С. 8–57.
3. *Хилько Д. В., Шикунов Ю. И.* Разработка инструментальной среды проектирования программного обеспечения для рекуррентно-потоковой модели вычислений // Четвертая школа молодых ученых ИПИ РАН, 2013. Сборник докладов. С. 65–77.
4. *Хилько Д. В., Степченков Ю. А.* Средства имитационного моделирования потоковой рекуррентной архитектуры (СИМПРА). Версия 2. Свидетельство о государственной регистрации программы для ЭВМ № 2014610123.
5. *Хилько Д. В., Степченков Ю. А., Шикунов Ю. И.* Средства имитационного моделирования многоядерной потоковой рекуррентной архитектуры // Многоядерные процессы, параллельное программирование, ПЛИС, системы обработки сигналов: сборник научных статей Всероссийской научно-практической конференции, Барнаул, 28 февраля 2014 г. – С. 58–69.

# Возможности параллельного программирования в пакетах MATLAB и Maple

А. М. Чернецов<sup>62</sup>

**Аннотация:** В работе рассмотрены средства параллельного программирования в современных математических пакетах. Для MATLAB и Maple кратко описываются имеющиеся средства и их аналоги в среде MPI.

**Ключевые слова:** параллельное программирование; MATLAB; Maple

## Введение

За последние годы при решении множества трудновычислимых задач стали применяться средства и возможности параллельного программирования. Широко известны модели программирования в общей и распределенной памяти [1], позднее к ним прибавились гибридные модели [2]. Однако все упомянутые средства относятся к достаточно низкоуровневому программированию, когда производятся значительная переделка исходного кода.

Немалое число математических расчетов выполняются не на алгоритмических языках (C/C++, Fortran), а в специализированных математических пакетах MATLAB, Maple, Mathematica, MathCad. Производители пакетов тоже задумались о реализации возможностей параллелизма в своих продуктах.

К 2015 году сложилась следующая картина. Изначально возможности параллельного программирования в пакете Mathematica появились в 2001 г. [3]. Для их обеспечения был выпущен дополнительный Parallel Computing Toolkit (отдельная ли-

---

<sup>62</sup> ФИЦ ИУ РАН (ВЦ им. А. А. Дородницына РАН); 1983 г.р.; к.т.н.; an@ccas.ru

цензия). После выхода версии 7.0 в 2008 г. для работы на локальном компьютере лицензия уже не требуется.

В пакете MATLAB возможности параллельного программирования появились в 2005 г. В 2009 г. о реализации средств параллельного программирования было заявлено в Maple. В среде Maple 13 было реализовано программирование в модели общей памяти. Однако средства именно параллельного программирования появились только в выпуске Maple 15 в 2011 г. Следует отметить, что независимые решения предлагались и ранее, см., например, Distributed Maple [4].

Пакет MathCad чрезвычайно долгое время никак не развивался в этом направлении. Только в MathCad Prime 2.0 (2012 г.) появились некоторые изменения – параллелизм как использование средств многоядерных процессоров обеспечивается за счет использования библиотеки Intel MKL. Далее не наблюдается никакого развития.

В данном докладе рассматриваются средства параллельного программирования в пакетах MATLAB и Maple.

## **1. Средства параллельного программирования в MATLAB**

### **1.1. Историческое развитие**

До момента реализации функционала параллельного программирования производителем существовали различные подходы к решению [5]. После обеспечения возможностей параллельного программирования производителем (компанией MathWorks) появились два вида параллелизма: distributed computing (распределенные вычисления), когда фактически обеспечивается параллелизм по данным [1], и собственно parallel computing. Для обеспечения возможностей параллелизма в пакете был разработан тулбокс Parallel Computing Toolbox (лицензируется отдельно).

Параллельное программирование в MATLAB основано на парадигме модели передачи сообщений. Последний на момент публикации выпуск R2015b использует MPI2 v.

1.4.1р1. Кроме того, допускается работа с собственной реализацией MPI, удовлетворяющей определенным условиям [6]. Для конечного пользователя есть два режима работы: интерактивный режим `pmode` [5], и, непосредственно, параллельное программирование в модели передачи сообщений (модели распределенной памяти).

## 1.2. Программирование в модели распределенной памяти в MATLAB

В Таблице 1 приведены основные примитивы параллельного программирования в MATLAB.

**Таблица 1.** Команды (переменные) MATLAB-и их соответствие в MPI

Команда MATLAB	Команда MPI
<code>Numlabs</code>	<code>MPI_Comm_size(size)</code>
<code>Labindex</code>	<code>MPI_Comm_rank(rank, MPI_COMM_WORLD)</code>
<code>LabBarrier()</code>	<code>MPI_Barrier(MPI_COMM_WORLD)</code>
<code>Shared_data=LabBroadcast(root,buffer)</code>	<code>MPI_Bcast(void* buffer, int count, MPI_Datatype datatype, int root, MPI_COMM_WORLD)</code>
<code>LabSend(buf,dest)</code> <code>MPI_Send(void *buf, int count, LabSend(data,dest,tag)</code>	<code>MPI_Send(void *buf, int count, int tag, MPI_COMM_WORLD)</code>
<code>data=LabReceive(source,tag)</code> <code>MPI_Recv(void*buf,int count, data=LabReceive('any',tag)</code>	<code>MPI_Recv(void *buf, int count, MPI_COMM_WORLD, MPI_Status *status)</code>
<code>is_data_available=LabProbe(source,tag)</code>	<code>MPI_Probe(int source, int tag, MPI_COMM_WORLD, MPI_Status *status)</code>

Как видно из таблицы, пакет реализует базовый набор директив для параллельного программирования.

Достаточно часто встречаются типовые ситуации, имеющие понятный способ решения. Для некоторых таких случаев

предусмотрены штатные средства. Рассмотрим два из них: векторизацию циклов (директива `parfor`) и частный случай SIMD (в классификации Флинна [1] – Single Program Multiple Data (директива `spmd`)).

Цикл `parfor` - параллельная модификация стандартного цикла `for`. Оператор `parfor` полезен в случаях, когда нужно провести много однотипных расчетов, например, при моделировании методом Монте-Карло. Итерации цикла считаются независимо в произвольном порядке рабочими процессами. Понятно, что оператор `parfor` нельзя применять в случаях, когда имеется зависимость по данным между итерациями.

Операторы, выполняемые внутри блока `spmd`, выполняются одновременно во всех рабочих процессах. Синтаксис команды:

```
spmd [(m,n)]  
<statements>  
end spmd
```

Здесь опциональные параметры `m` и `n` – минимальное и максимальное число рабочих процессов, соответственно.

Помимо описанных способов, начиная с MATLAB R2007b, в системе есть автоматическое распараллеливание функций различных тулбоксов (ядро, Image Processing, Statistics, Optimization Toolbox, и т.д.). Система производит автоматическое распределение расчетов на все доступные мощности (например, на все 8 ядер процессора).

Кроме того, MATLAB с выпуска R2010b поддерживает параллельные вычисления на графических процессорах (GPU) NVIDIA – программирование в модели CUDA. Соответственно, есть директивы для копирования данных из основной памяти в GPU и обратно, директивы для выполнения операций на GPU.

MATLAB поддерживает выполнение параллельных задач на кластерных системах с использованием функционала MATLAB Distributed Computing Engine (лицензируется отдельно). В его рамках возможен запуск параллельных задач в СУПЗ Microsoft, PBS, Torque, LSF, Condor [7].

## 2. Средства параллельного программирования в Maple

### 2.1. Программирование в модели общей памяти

В версии Maple 13 появилась Task Programming Model, которая позволяет работать в модели общей памяти на высоком уровне абстракции. На более низком уровне Task Programming Model реализуется как работа с нитями (threads). Разработчик рекомендует конечным пользователям не работать на этом уровне.

При выполнении программы Maple создаются подзадачи (Tasks), которые средой распределяются на все доступные ресурсы процессоров. Каждая подзадача – это вызов функции, который выполняется отдельной нитью.

### 2.2. Программирование в модели распределенной памяти

В развитие, в версии Maple 15 появились средства именно параллельного программирования, которые появились в дополнительном пакете Grid Package (Toolbox). Так же, как и в MATLAB, пакет позволяет работать с различными реализациями MPI. По умолчанию в среде Windows используется Microsoft MPI, включенный в Windows HPC Server, а в среде Linux – MPICH2.

В Таблице 2 приведены основные примитивы параллельного программирования в Maple. Более подробно с ними можно ознакомиться в документации на продукт.

**Таблица 2.** Функции Maple-и их соответствие в MPI

<b>Функция Maple</b>	<b>Команда MPI</b>
NumNodes()	MPI_Comm_size(size)
MyNode()	MPI_Comm_rank(rank, MPI_COMM_WORLD)
Barrier()	MPI_Barrier(MPI_COMM_WORLD)

Set(node,v1, v2, ...)	int MPI_Scatter(void* sendbuf, int sendcount, MPI_Datatype sendtype, void* recvbuf, int recvcnt, MPI_Datatype recvtype, int root, MPI_COMM_WORLD)
Send( node, msg)	MPI_Send(void *buf, int count, int tag, MPI_COMM_WORLD)
Receive() Receive(node) - получить от конкретного узла	MPI_Recv(void *buf, int count, MPI_COMM_WORLD, MPI_Status *status)

Как видно из таблицы, пакет реализует базовый набор директив для параллельного программирования, при этом он отличается по функционалу от MATLAB. Также отметим, что весь функционал уже лицензирован в рамках пакета, и не требуется приобретения дополнительных лицензий.

Для указания на то, что вычисления необходимо производить параллельно, служит директива Launch.

Пакет Maple, также как и MATLAB, позволяет производить параллельные вычисления на GPU. Maple поддерживает выполнение параллельных задач на кластерных системах с использованием функционала тулбокса Grid. В его рамках возможен запуск параллельных задач в СУПЗ Microsoft, PBS, Torque. В этом случае уже требуется дополнительная лицензия на удаленных клиентах-узлах.

## Выводы

В заключение хотелось бы сказать, что средства параллельного программирования в математических пакетах постоянно развиваются. В зависимости от того, в каком пакете происходит работа, доступен несколько отличающийся функционал, однако любая задача может быть решена как в Maple, так и в MATLAB.

## Список литературы

1. *Воеводин В.В., Воеводин В.В.* Параллельные вычисления. –

Санкт-Петербург: ВHV-Петербург, 2002. – 608 с.

2. *Эндрюс Г.Р.* Основы многопоточного, параллельного и распределенного программирования. М.: Вильямс, 2003. 512 с.
3. *Mäder R.* Parallel Computing with Mathematica //Proceedings of the 4th International Workshop on Computer Algebra in Scientific Computing, Konstanz, Sept.22-26, 2001. P. 399.
4. *Schreiner W., Mittermaier C., Bosa K.* Distributed Maple: parallel computer algebra in networked environments //Journal of Symbolic Computation, 2003. Vol.35. № 3. P. 305–347.
5. *Оленев Н. Н., Печенкин Р. В., Чернецов А. М.* Параллельное программирование в MATLAB. – М.: ВЦ РАН, 2007. 117 с.
6. *Чернецов А.М.* Использование сторонних библиотек MPI для параллельного программирования в MATLAB // Радиоэлектроника, электротехника и энергетика. 14-я межд. научно-техническая конф. студентов и аспирантов в 3-х т., т. 1 - М.: МЭИ, 2008. С. 309–310.
7. *Чернецов А.М.* Использование средств MATLAB для организации распределенной обработки // Труды межд. конф. «Информатизация инженерного образования», 10-11 апреля 2012 г. - М.: МЭИ, 2012. С. 127–130.

# О методах оценки цветопередачи источников света путём анализа спектра с использованием языка программирования Python

А.А. Шинов<sup>63</sup>

**Аннотация:** Данная работа рассматривает существующие методы и предлагает новый для оценки качества передачи цветов источником света. Проводится сравнение методов на примере реального исследования. Для предложенного метода оценки используются средства анализа данных в языке программирования Python, анализ спектра проводится в диапазоне 380 ÷ 740 нм (видимое человеческим глазом излучение).

**Ключевые слова:** цветопередача; Python; анализ спектра

## Введение

Одним из важных критериев оценки источников света является критерий цветопередачи, определяющий, насколько цвета, видимые при освещении объектов оцениваемым источником, близки к цветам, видимым при освещении тех же объектов эталонным источником. Проблема оценки цветопередачи является актуальной и не решённой на достаточном для ряда повседневных (и не только) задач уровне.

В данной работе будет предложен метод оценки цветопередачи, основанный на анализе спектра источника света и дающий более точные результаты, чем существующие методы.

## 1. Методы оценки цветопередачи

### 1.1. Существующие методы оценки цветопередачи

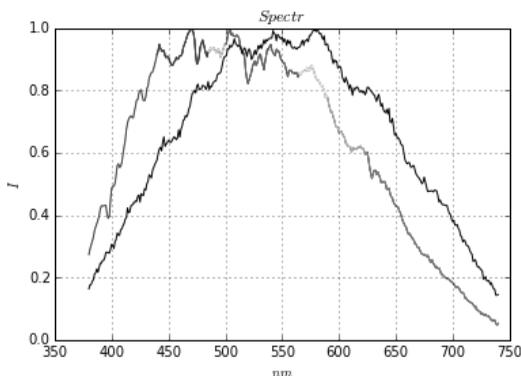
Показателем, характеризующим качество передачи

---

<sup>63</sup> МФТИ; 1994 г.р.; alexandershinov@gmail.com

цветов источником света, может служить коэффициент (или индекс) цветопередачи. Существует несколько стандартных методов его оценки.

Первый метод заключается в оценке в среднем по восьми эталонным цветам. Цветопередача определяется отклонениями от эталонов в цветовом пространстве CIE XYZ [1, 2]. Невысокую точность этого метода наглядно демонстрирует то, что лампа накаливания согласно этому методу обладает идеальной цветопередачей, что означает полное подобие спектров ламп накаливания с естественным светом, однако это не соответствует действительности, что наглядно иллюстрируется на Рис. 1.



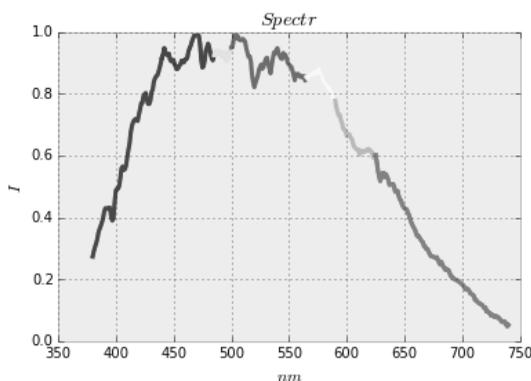
**Рис. 1.** Сравнение спектра лампы накаливания (однородная чёрная кривая) с солнечным спектром

Наряду с ним существует и более точный метод оценки по шкале CQS (Color Quality Scale), разрабатываемой американским институтом NIST (Государственный институт стандартов и технологий). Его шкала корректно оценивает спектр светодиодов, при этом при оценке цветовоспроизведения высокотемпературных ламп и ламп накаливания появляются существенные погрешности [4].

## 1.2. Предлагаемый метод оценки цветопередачи

Для реализации предлагаемого ниже метода использо-

вался язык программирования Python версии 2.7 с подключенными библиотеками NumPy (`import numpy as np`) и Matplotlib (`import matplotlib.pyplot as plt`). Для удобства работы в режиме реального времени использовалась среда IPython [3]. Предлагаемый в данной работе метод оценки цветопередачи позиционирует результирующий коэффициент как процентное значение качества передачи цвета относительно эталонного (естественного, солнечного) источника. При помощи спектрометра был получен спектр солнечного света, изображённый на Рис. 2.



**Рис. 2.** Солнечный спектр

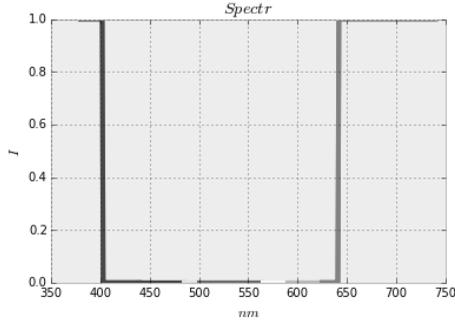
Путём определения максимально возможного отклонения от солнечного спектра на диапазоне 380-740 нм была построена функция нулевой цветопередачи. Она изображена на Рис. 3.

Введем показатель цветопередачи (1) как отношение площади между изучаемым спектром и эталоном к площади между эталоном и функцией нулевой цветопередачи ( $\text{solar}(x)$  – функция солнечного спектра,  $\text{nully}(x)$  – функция нулевой цветопередачи,  $f(x)$  – функция оцениваемого спектра). На Рис. 4 показаны соответствующие зависимости и площади.

```

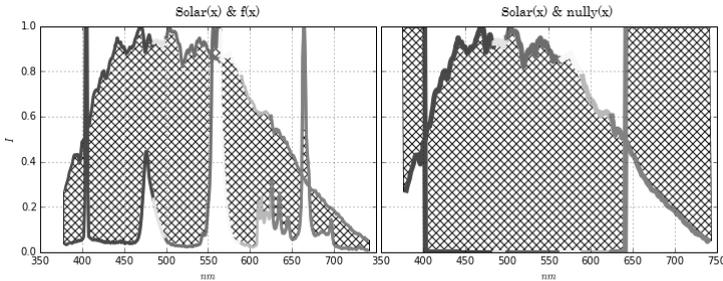
In [1]: %run D:\kfs.py
In [2]: nully = np.where(solar_y>1-solar_y, 0, 1)
In [3]: DrawSpectr(nully)

```



**Рис. 3.** Функция нулевой цветопередачи

$$K_c = \frac{\int_{380}^{740} |solar(x) - f(x)| dx}{\int_{380}^{740} |solar(x) - nully(x)| dx} \quad (1)$$



**Рис. 4.** Иллюстрация применения формулы (1):  
слева – числитель, справа – знаменатель

На Рис. 5 изображён код функции, вычисляющий показатель цветопередачи на языке программирования Python. Средства библиотеки NumPy [3] позволяют быстро решить поставленную вычислительную задачу, благодаря удобному инструментарию для работы с массивами. Метод where позволяет по определённому логическому условию без использования стандартного механизма циклов подставить

в каждый отдельный элемент массива нужное значение, зависящее от выполнения определенного условия. Метод `sum` складывает все значения элементов массива.

```
def GetKc(y):
    max_a = np.where(solar_y>1-solar_y, np.abs(solar_y), np.abs(1-solar_y)).sum()
    a = (np.abs(y - solar_y)).sum()
    print 'Kc = ', 1-a/max_a
```

**Рис. 5.** Функция вычисления показателя цветопередачи

Таким образом, разработанный выше подход может быть применен к оценке правильности цветопередачи освещаемых объектов.

## 2. Результаты оценки нескольких реальных источников

С использованием разработанного метода были оценены 12 различных источников, также было проведено сравнение со стандартным методом оценивания по шкале CRI. Результаты представлены ниже в Таблице 1. В столбце `Kc` – результаты по предлагаемому методу, в столбце `CRI` – стандартная оценка по шкале CRI, где класс 1A является наивысшей оценкой.

**Таблица 1.** Исследованные источники света

№	Тип источника	Kc	CRI
1	ЛМЛ	42%	2B
2	Накаливания	78%	1A
3	Накаливания	80%	1A
4	Накаливания	80%	1A
5	Накаливания	85%	1A
6	Накаливания	74%	1A
7	Накаливания	78%	1A
8	На светодиодах	59%	2A
9	ЛМЛ	35%	2B
10	ЛМЛ	35%	2B
11	ЛМЛ	42%	2B
12	Галогенная	71%	1A

Таким образом, отсутствуют порядковые конфликты (оба метода одинаково расставляют источники света по возрастанию/убыванию качества цветопередачи). При этом предложенный в работе метод не оценивает ни один искусственный источник как идеальный. Если предположить, что появится искусственный источник света, полностью соответствующий естественному по спектру излучения, на шкале CRI он будет соответствовать уровню лампы накаливания или галогенной лампы. В то время как предложенный метод даст оценку нового источника как более качественного по параметру цветопередачи.

## **Выводы**

Разработанный метод находится в согласии с наиболее распространённым методом оценки цветопередачи по шкале CRI. Его отличительной особенностью является более высокая точность при оценке качественных по цветопередаче источников света по сравнению с другими стандартными методами.

Метод реализован с использованием инструментария языка программирования Python и позволяет в режиме реального времени анализировать характеристики параметров спектра для получения более качественных выводов о цветопередаче.

## **Список литературы**

1. *Джадд Д., Вышецкий Г.* Цвет в науке и технике. – М.: Мир, 1978. – 592 с.
2. *Гуревич М. М.* Цвет и его измерение. – М.: Издательство Академии наук СССР, 1950. – 269 с.
3. *Маккинли У.* Python и анализ данных. – М.: ДМК Пресс, 2015. – 482 с.
4. *Pousset N., Obein G., Razet A.* Lighting Quality and Energy Efficiency // CIE 2010. – Vienna., 2010. – 7 p.

# Метод многослойного сетевого оператора в задаче синтеза системы управления группой роботов\*

Е. Ю. Шмалько<sup>64</sup>, А. И. Дивеев<sup>65</sup>

**Аннотация:** При решении задачи синтеза системы управления группой роботов существенно увеличивается количество аргументов искомого математического выражения. Это приводит к увеличению размерности матрицы сетевого оператора. Для преодоления указанной проблемы предлагается использовать метод многослойного сетевого оператора. В работе представлено описание метода, его основные особенности и преимущества.

**Ключевые слова:** многослойный сетевой оператор; группа роботов; численный синтез управления

## Введение

Потенциальные преимущества использования групп роботов вместо роботов-одиночек часто обсуждаются в современной литературе [1, 2]. Например, общая стоимость владения может быть уменьшена за счет использования множества простых и дешевых роботов вместо одного сложного и дорогостоящего робота. Некоторые задачи могут быть выполнены более эффективно группой роботов за счет разбиения основной задачи на подзадачи, которые можно выполнить параллельно. Сами условия выполнения задачи могут оказаться настолько сложными, что для реализации поставленной цели потребуются значительные способности, которые просто невозможно реализовать в одном роботе, но реально сделать группой роботов. И, наконец, считается, что использование группы роботов увеличивает робастность системы за счет присущего такой системе

---

\*Работа выполнена при поддержке гранта Президента РФ МК-6277.2015.8.

<sup>64</sup> ФИЦ ИУ РАН (ВЦ им. А. А. Дородницына РАН); 1982 г.р.; к.т.н.; e.shmalko@gmail.com

<sup>65</sup> ФИЦ ИУ РАН (ВЦ им. А. А. Дородницына РАН); 1954 г.р.; д.т.н.; aidiveev@mail.ru

параллелизма и избыточности. Также использование группы роботов позволяет минимизировать риски, связанные с невыполнением поставленной задачи, которые могут возникнуть за счет выхода из строя одного или нескольких роботов.

Среди общих подходов к созданию систем управления группами роботов можно выделить определенные классы.

К первому классу с равноправным управлением можно отнести такие группы роботов, у которых управление равноправное, группа состоит из одинаковых роботов и все роботы обладают полную информацию друг о друге.

Ко второй группе с ранжированным управлением можно отнести задачи, где роботы ранжированы по знанию информации о других роботах. Примером такого подхода являются группы, организованные по принципу подчинения, где роботы высшего ранга имеют максимально полную информацию, а роботы низшего ранга могут вообще не иметь информации о внешней среде, они полностью подчиняются командам роботов высшего ранга.

К третьему смешанному классу можно отнести группы роботов, где роботы равноправны, но обладают информацией только о себе и о некоторых других роботах, например, расположенных по соседству. К этому же классу относится и подкласс, где роботы равноправны, но с переменной информацией, например, при движении роботов каждый из них знает информацию о положении только тех роботов, которые расположены в данный момент в его окрестности.

Всем трем выделяемым классам управления группой роботов присуща общая особенность, связанная с наличием в задаче динамических фазовых ограничений, обеспечивающих отсутствие столкновений роботов друг с другом во время движения. Проверка выполнения данных условий на каждом шаге управления значительно увеличивает вычислительную нагрузку на систему управления. Так в случае управления группой, состоящей из двух роботов, проверять необходимо 4 условия. В случае с тремя роботами таких условий будет 6. И

далее при увеличении числа роботов число условий будет возрастать согласно  $n(n-1)$ .

На сегодняшний день для синтеза системы управления необходимо найти управление в виде многомерной нелинейной функции от состояния объекта управления. Искомая функция может быть негладкой и разрывной, т.е. иметь разрывы первого рода. Нахождение такой функции и ее реализация на бортовых компьютерах объектов управления позволяют создать систему управления, которая сможет перевести объекты из начального состояния на терминальное многообразие с оптимальным значением критериев качества.

## **1. Метод многослойного сетевого оператора**

Для решения задачи используем численный метод сетевого оператора [3]. Метод относится к классу методов символьной регрессии [4]. Суть метода заключается в том, что код математического выражения описывается в виде целочисленной матрицы, в ячейках которой хранятся номера переменных, параметров, унарных и бинарных операций. Поиск математического выражения, удовлетворяющего целям и критериям, осуществляется с помощью эволюционного алгоритма на основе принципа вариации базисного решения.

При увеличении числа роботов увеличивается количество входов и выходов сетевого оператора и соответственно значительно возрастает размерность матрицы сетевого оператора. Для преодоления указанных трудностей предложена новая концепция многослойного сетевого оператора. Многослойный сетевой оператор представляет собой связанную последовательность сетевых операторов малой размерности, когда значения на выходе одного сетевого оператора могут являться входами для другого. Представленная новая концепция многослойного сетевого оператора позволяет сколь угодно большой сетевой оператор представлять в виде связанного множества сетевых операторов малой размерности. Разработанная технология поз-

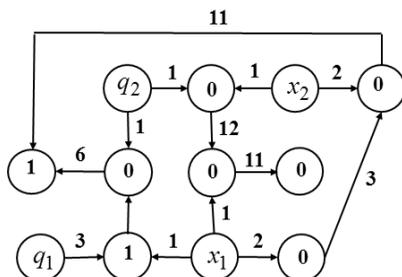
воляет решать задачу синтеза системы управления группой роботов без ограничения на их количество.

Концепцию организации вычислений на основе многослойного сетевого оператора рассмотрим на примере.

Пусть имеется функция

$$\begin{aligned} y_1 &= \cos(x_1 + \sin(x_2 + q_2)), \\ y_2 &= e^{-q_1 x_1 + q_2} \cos(x_1^2 - x_2^2). \end{aligned} \quad (1)$$

В ориентированном графе сетевого оператора элементы множеств переменных и параметров связаны с узлами-источниками графа, бинарные операции связаны с остальными узлами графа, унарные операции связаны с дугами графа. Граф сетевого оператора выражения (1) представлен на Рис.1. Номера бинарных и унарных операций соответствуют приложению монографии [3].



**Рис. 1.** Сетевой оператор выражения

Представим (1) в матричном виде многослойного сетевого оператора со следующими параметрами: количество слоев  $N = 6$ , размерность каждого слоя  $L \times L$ ,  $L = 6$ , количество входов на каждом слое  $I = 2$ . Матрицы каждого слоя сетевого оператора будут иметь вид

$$\Psi_1 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 12 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \Psi_2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 11 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\Psi_3 = \begin{bmatrix} 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 11 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \Psi_4 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\Psi_5 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \Psi_6 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Вводим нулевой слой или вектор входов многослойного сетевого оператора. В рассматриваемом примере размерность вектора входов соответствует количеству узлов-источников графа сетевого оператора (см. Рис. 1).

$$\mathbf{z}^0 = [x_1 \ x_2 \ q_1 \ q_2]^T.$$

В общем случае вектор входов включает в себя компоненты вектора состояния, компоненты вектора наблюдения, параметры, которые необходимо оптимизировать, константы, неопределенные параметры.

Для описания связей между слоями вводим матрицу связей размерности  $N \times (2I)$ . Каждая строка матрицы соответствует номеру слоя. Каждый нечетный столбец матрицы указывает на номер входа данного слоя, откуда приходит входная величина, а значение четного столбца указывает на номер строки матрицы, от которой приходит входная величина. Если указан номер слоя 0, то входная величина берется из вектора входов, соответственно, значение элемента четного столбца указывает на номер компоненты вектора входов.

Матрица связей в рассматриваемом примере имеет вид

$$\Sigma = \begin{bmatrix} 0 & 2 & 0 & 4 \\ 0 & 1 & 1 & 6 \\ 0 & 1 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 4 & 4 & 6 \\ 3 & 6 & 5 & 6 \end{bmatrix}.$$

Матрица выходов имеет размерность  $2 \times M$ , где  $M$  – число кодируемых математических выражений или число выходов. В рассматриваемом случае  $M = 2$  и

$$\Omega = \begin{bmatrix} 2 & 6 \\ 6 & 6 \end{bmatrix},$$

где в верхней строке указывается номер слоя, а в нижней – соответствующий номер выхода.

Для вычисления математического выражения используем следующий алгоритм

$$z_{k,r} = \begin{cases} z_0^0 \sigma_{k,2r}, & \text{если } \sigma_{k,2r-1} = 0 \\ z_{\sigma_{k,2r-1}, \sigma_{k,2r}} & \text{иначе} \end{cases},$$

$$z_{k,j} \leftarrow \begin{cases} \chi_{\psi_{k,j,i}}(z_{k,j}, \rho_{\psi_{k,i,j}}(z_{k,i})) & \text{если } \psi_{k,i,j} \neq 0 \\ z_{k,j} & \text{иначе} \end{cases},$$

где  $\Psi_k = [\psi_{k,i,j}]$ ,  $\Sigma = [\sigma_{k,m}]$ ,  $m = \overline{1, 2I}$ ,  $k$  – номер слоя,  $k = \overline{1, N}$ ,  $r$  – номер входа,  $r = \overline{1, I}$ ,  $i$  – номер строки матрицы сетевого оператора,  $i = \overline{1, L}$ ,  $j$  – номер столбца сетевого оператора,  $j = \overline{i+1, L-2}$ ,  $\rho_d(z)$  – унарная операция под номером  $d$ ,  $\chi_g(z_1, z_2)$  – бинарная операция под номером  $g$ .

Значение кодируемой функции определяем из соотношения

$$y_p = z_{\omega_{1,p}, \omega_{2,p}}, \quad p = \overline{1, M},$$

где  $\Omega = [\omega_{l,t}]$ ,  $l = \overline{1, M}$ ,  $t = 1, 2$ .

Программа для работы с многослойным сетевым оператором в настоящий период отлажена в среде программирования Lazarus и зарегистрирована в фонде программ [5].

## Выводы

Сетевой оператор представляет собой универсальную структуру для программного кодирования различных математических выражений, которые могут описывать не только

управление, но и правые части дифференциальных уравнений, описывающих объекты, терминальные условия, функционалы и т.д. И в данном ключе структура многослойного сетевого оператора является более общей формой представления, поскольку включает в себя вектор входов, состоящий из вектора состояния, вектора наблюдения, параметров, которые необходимо оптимизировать, константы и неопределенные параметры. И в зависимости от структуры, которую мы описываем, достаточно указывать только номера компонент из общего вектора входов и связи между слоями.

## Список литературы

1. *Parker L.* Alliance: An architecture for fault-tolerant multi-robot cooperation // Transactions on Robotics and Automation, 14(2), pp. 220–240, Apr. 1998.
2. *Arkin R., Balch T.* Cooperative multiagent robotic systems / D. Kortenkamp, R. P. Bonasso, R. Murphy (editors), Artificial Intelligence and Mobile Robots: Case Studies of Successful Robot Systems, MIT/AAAI Press, Cambridge, MA, 1998. P.277–296.
3. *Дивеев А. И.* Метод сетевого оператора. – М.: ВЦ РАН, 2010. – 178 с.
4. *Zelinka I., Oplatkova Z., Nolle L.* Analytic programming - symbolic regression by means of arbitrary evolutionary algorithms // Int. J. Simulation Syst. Sci. Technol. Vol. 6 No 9. PP. 44–56.
5. *Дивеев А. И., Шмалько Е. Ю.* Программа расчета математического выражения с помощью многослойного сетевого оператора// Программа ЭВМ № 2015614591 в реестре программ для ЭВМ 15.07.2015.

# Анализ системы поллинга с пороговым управлением\*

Ш.А. Этезов<sup>66</sup>

**Аннотация:** Статья содержит описание исследования моделей управления перегрузками SIP-серверов в виде системы поллинга. Рассмотренная модель подразумевает размещение всех поступающих на SIP-сервер сообщений в две очереди. Обслуживание очередей производится циклично в соответствии с различными дисциплинами. В рамках численного анализа было проведено сравнение полученных результатов для двух возможных дисциплин обслуживания: исчерпывающей и шлюзовой.

**Ключевые слова:** протокол SIP, контроль перегрузок, пороговое управление нагрузкой, система массового обслуживания, поллинг, исчерпывающая дисциплина, шлюзовая дисциплина

## Введение

Протокол инициирования сеансов связи (Session Initiation Protocol, SIP) [1] является одним из основных сигнальных протоколов голосовых сетей и обеспечивает установление соединений между их абонентами. Неотъемлемой частью схемы при использовании протокола SIP также являются SIP-серверы, обеспечивающие обмен сигнальными сообщениями между абонентами. В некоторых случаях при резком увеличении сигнального трафика на SIP-серверах возможны перегрузки.

Существующий механизм управления перегрузками в соответствии с [2] недостаточно проработан. Его использование в некоторых случаях может привести к еще большим нагрузкам на сеть. Поэтому активно ведутся разработки альтернативных

---

\*Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 15-07-03608

<sup>66</sup> РУДН; 1989 г.р.; setezov@gmail.com

механизмов, которые позволят обеспечивать заданное качество обслуживания в сетях передачи речи. Одним из таких механизмов может быть метод порогового контроля перегрузок, основанный на гистерезисном управлении нагрузкой, предложенный в [3, 4], в сочетании с системами поллинга.

Предлагаемый механизм подразумевает разделение всех сообщений, поступающих на SIP-сервер на неприоритетные (сообщения Invite) и приоритетные (сообщения nonInvite). Для каждого из двух типов сообщений на SIP-сервере выделяется буфер определенного размера. Обслуживающий прибор (процессор) осуществляет циклический опрос двух очередей в соответствии с определенной дисциплиной. В работе рассматривались исчерпывающая и шлюзовая дисциплины обслуживания. В случае использования исчерпывающей дисциплины прибор не покидает очереди, пока не обработает все находящиеся в ней сообщения. В случае использования шлюзовой дисциплины прибор обслуживает за один цикл только те сообщения, которые были в очереди в момент его переключения. После обслуживания данных сообщений прибор переключается на другую очередь.

Анализ моделей проведён при помощи аппарата марковских случайных процессов. Для моделей с исчерпывающей и шлюзовой дисциплинами обслуживания были получены формулы для расчета вероятностно-временных характеристик, после чего было проведено сравнение полученных результатов.

## 1. Описание механизма

### 1.1. Математическая модель

Рассматривается математическая модель функционирования SIP-сервера в виде системы массового обслуживания (СМО) с двумя очередями и одним обслуживающим прибором. Очереди имеют конечную емкость  $r_1 < \infty$  и  $r_2 < \infty$ ,  $r_1 + r_2 = R$ . Поступающие на систему потоки заявок являются пуассоновскими с интенсивностями  $\lambda_1$  и  $\lambda_2$ . Длительности обслуживания

заявок в каждой из очередей имеют экспоненциальное распределение с параметрами  $\mu_1$  и  $\mu_2$  соответственно. Будем считать, что первая очередь содержит приоритетные заявки (сообщения nonInvite) и введем в этой очереди порог  $0 < L < r_1$ . После достижения данного порога снижается интенсивность поступления сообщений Invite в неприоритетную вторую очередь до величины  $\lambda'_2 < \lambda_2$ . Время переключения между очередями также считаем случайным, распределенным в соответствии с экспоненциальным законом с параметрами  $s_1$  в случае переключения со второй очереди на первую и  $s_2$  в случае переключения с первой очереди на вторую. Полагается, что обслуживаемая прибором заявка сохраняет за собой место в очереди до момента ухода из системы. Схематично рассматриваемая СМО представлена на Рис. 1.

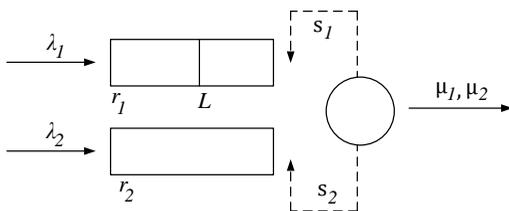


Рис. 1. СМО с поллингом и пороговым управлением

## 1.2. Исчерпывающая дисциплина

Определим случайный процесс  $\eta(t) = (I(t), n_1(t), n_2(t))$ , где  $I(t)$  – номер очереди, в которой находится прибор, и  $n_i(t)$  – число заявок в  $i$ -й очереди в момент  $t > 0$ . Случайный процесс является марковским над пространством состояний

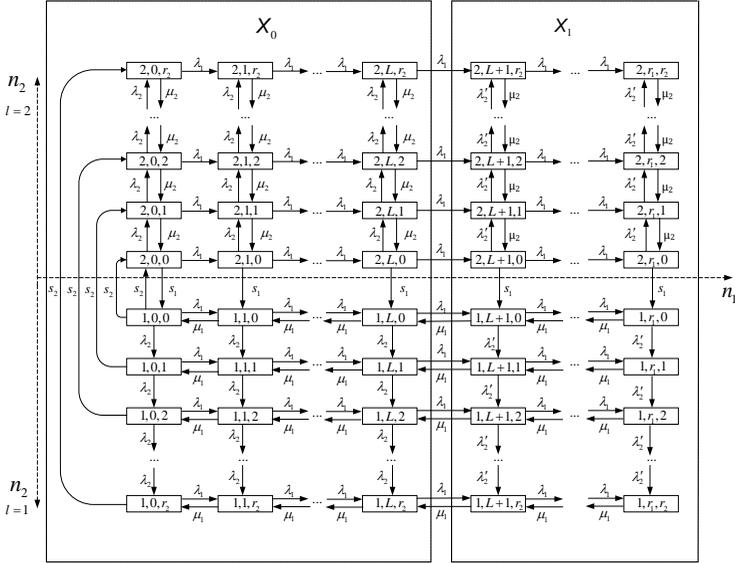
$$X = \{(I, n_1, n_2) \mid I \in \{1, 2\}, n_1 \in \{0, \dots, r_1\}, n_2 \in \{0, \dots, r_2\}\}, \quad (5)$$

причем  $|X| = 2(r_1 + 1)(r_2 + 1)$ .

$$X_0 = \{(I, n_1, n_2) \mid I \in \{1, 2\}, n_1 \in \{0, \dots, L\}, n_2 \in \{0, \dots, r_2\}\}, \quad (6)$$

$$X_1 = \{(l, n_1, n_2) \mid l \in \{1, 2\}, n_1 \in \{L+1, \dots, r_1\}, n_2 \in \{0, \dots, r_2\}\} \quad (7)$$

являются подмножествами нормальной нагрузки и перегрузки соответственно. Диаграмма интенсивностей переходов между состояниями определенного марковского процесса представлена на Рис. 2.



**Рис. 2.** Диаграмма интенсивностей переходов для исчерпывающей дисциплины обслуживания

### 1.3. Шлюзовая дисциплина

Определим случайный процесс  $\eta(t) = (l(t), n_1(t), n_2(t), m(t))$ , где  $l(t)$  – номер очереди, в которой находится прибор,  $n_i(t)$  – число заявок в  $i$ -й очереди,  $m(t)$  – число заявок в очереди с номером  $l(t)$ , которое осталось обслужить прибору в текущем цикле, в момент времени  $t > 0$ . Случайный процесс является марковским над пространством состояний

$$X = \left\{ (l, n_1, n_2, m) : m \in \{0, \dots, \max(r_1, r_2)\}; n_i \in \{m, \dots, r_i\}, \right. \\ \left. n_{(l+1) \bmod 2} \in \{0, \dots, r_{(l+1) \bmod 2}\}; l = 1, 2 \right\}, \quad (8)$$

причем  $|X| = \frac{1}{2}(r_2 + 1)(r_1 + 1)(r_1 + r_2 + 4)$ .

$$X_0 = \left\{ (l, n_1, n_2, m) : l \in \{1; 2\}; m \in \{0, \dots, r_l\}, n_l = \{m, \dots, r_l\}, \right. \\ \left. n_{(l+1) \bmod 2} \in \{0, \dots, r_{(l+1) \bmod 2}\}, n_1 \leq L \right\}, \quad (9)$$

$$X_1 = \left\{ (l, n_1, n_2, m) : l \in \{1; 2\}; m \in \{0, \dots, r_l\}, n_l = \{m, \dots, r_l\}, \right. \\ \left. n_{(l+1) \bmod 2} \in \{0, \dots, r_{(l+1) \bmod 2}\}, n_1 > L \right\} \quad (10)$$

являются подмножествами нормальной нагрузки и перегрузки соответственно.

## 2. Численный анализ

Для моделей с обеими дисциплинами обслуживания матрица интенсивностей переходов была приведена к блочному-трехдиагональному виду, что позволило численно решить систему уравнений равновесия (СУР) и найти стационарные вероятности. Для численных расчетов использовались следующие исходные данные:  $\lambda_1 = 6\lambda_2$ ,  $R = 20$ ,  $r_1 = r_2 = 10$ ,  $L = 5$ ,  $\lambda_2' = \lambda_2 / 2$ ,  $\mu_1^{-1} = 4$  мс,  $\mu_2^{-1} = 10$  мс,  $s_1^{-1} = s_2^{-1} = 10$  мс. На Рис. 3 представлено сравнение среднего числа заявок в очередях для моделей с исчерпывающей и шлюзовой дисциплинами обслуживания.

## Выводы

В работе исследован один из возможных механизмов управления перегрузками SIP-сервера в виде модели СМО с пороговым управлением и поллинговой системой обслуживания. В качестве поллинговых систем рассматривались исчерпывающая и шлюзовая дисциплины обслуживания.

Проведенный численный анализ показал некоторые преимущества шлюзовой дисциплины обслуживания заявок перед исчерпывающей дисциплиной с точки зрения среднего числа приоритетных сообщений в СМО и вероятности нахождения системы в состояниях нормальной нагрузки.

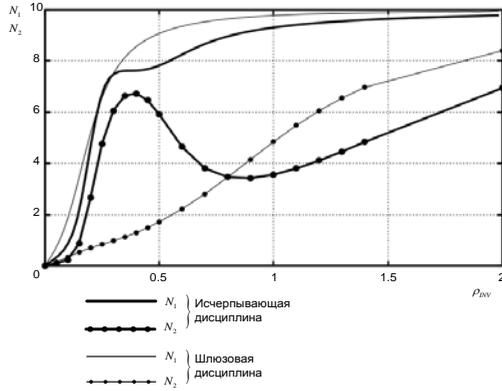


Рис. 3. Среднее число заявок в очередях

## Список литературы

1. *Rosenberg J., Schulzrinne H., Camarillo G. et al.* SIP: Session Initiation Protocol // RFC 3261. — June 2002.
2. *Rosenberg J.* Requirements for Management of Overload in the Session Initiation Protocol // RFC 5390. December 2008.
3. *Abaev P., Gaidamaka Yu., and Samouylov K.* Modeling of Hysteretic Signalling Load Control in Next Generation Networks // Proceedings 12th International Conference, NEW2AN 2012, and 5th Conference, ruSMART 2012, St. Petersburg, Russia, August 27-29, 2012. — LNCS, 2012. — Vol. 7469. P. 440-452.
4. *Abaev P., Gaidamaka Yu., and Samouylov K.* Queuing Model for Loss-Based Overload Control in a SIP Server Using a Hysteretic Technique // Proceedings 12th International Conference, NEW2AN 2012, and 5th Conference, ruSMART 2012, St. Petersburg, Russia, August 27-29, 2012. — LNCS, 2012. Vol. 7469. — P. 371–378.

# Современное состояние алгебраической теории анализа изображений. Обзор\*

В. В. Яшина<sup>67</sup>

**Аннотация:** Основная задача обзора показать возможности и ограничения алгебраической теории анализа изображений. В настоящее время принято считать, что применение алгебраических методов, в частности различных видов алгебр изображений, приведет к прорыву в развитии математической теории анализа изображений. Опыт развития последней и ее использование для решения прикладных задач показывают, что при работе с изображениями необходимо решать задачи, возникающие в связи с тремя основными проблемами анализа изображений: а) описанием и представлением изображений; б) разработкой, исследованием и оптимизацией выбора математических методов и средств обработки изображений; в) программной и аппаратной реализацией математических методов. В обзоре рассматриваются фундаментальные основы для решения этих задач в рамках алгебраической теории анализа изображений.

**Ключевые слова:** алгебры изображений; модели изображений; Дескриптивный подход к анализу и пониманию изображений

## Введение

Автоматизация обработки, анализа, оценивания и понимания информации, представленной в виде изображений, является одной из критических прорывных проблем теоретической информатики. Развитие этого направления информатики имеет решающее значение для решения прикладных задач, в частности, для увеличения разнообразия классов разрешимых задач и существенного увеличения эффективности их решения.

---

\*Работа выполнена при поддержке РФФИ (грант № 14-01-00881).

<sup>67</sup> ФИЦ ИУ РАН (ВЦ им. А. А. Дородницына РАН); 1980 г.р.; к.ф.-м.н.; [werayashina@gmail.com](mailto:werayashina@gmail.com)

Анализ изображений использует широкий набор математических методов алгебры, геометрии, дискретной математики и математической логики, теории вероятностей и математической статистики, математического анализа, а также методы математической теории распознавания образов, теории обработки сигналов и методы физики, в том числе оптики.

После 2000 г. ведущим направлением в развитии математической теории анализа изображений стала «алгебраизация» обработки, анализа и распознавания изображений, заключающаяся в разработке и исследовании алгебр изображений (АИ).

## 1. «Алгебраизация» анализа изображений

Среди достижений в области «алгебраизации» отметим работы Г. Риттера[16], Ж. Серра[18], С. Стенберга[20], Г. Матерона [12]. Основной целью алгебраического подхода является построение теоретического аппарата, обеспечивающего представление изображений и преобразований над ними в виде алгебраических структур, позволяющих стандартизировать анализ и распознавание изображений и использовать методы из различных областей математики.

Общепринятым методом стандартизации в современной прикладной математике и информатике является построение и использование математических и имитационных моделей изучаемых объектов и процедур, используемых для их преобразования. Среди существенных достижений этого «дескриптивного» направления отметим работы А. Розенфелда[17], Т. Эванса[4], Р. Нарасимхана[14], Р. Кирша[11], А. Шоу[19], Г. Барроу[3], Ф. Эмлера[3], Р. Бэрстолла[3], С. Кейнеффа[10].

Развитие математической теории анализа изображений идёт по пути её алгебраизации на основе алгебраического подхода к задачам распознавания и классификации, который был развит Ю. И. Журавлёвым [2], и его специализации на случай представления исходной информации в виде изображений – Дескриптивный подхода к анализу и пониманию изображений (ДПИ), предложенный и развиваемый И. Б. Гуревичем [5], а

также разработки теории алгебр изображений [7].

Ю. И. Журавлев [2] предложил т.н. «Алгебраический подход к задачам распознавания и классификации», в котором определил методы формализации описания эвристических алгоритмов распознавания образов и ввел универсальную структуру алгоритма распознавания.

Основной целью ДП [5-9] является разработка математического алгебраического аппарата для моделирования изображений и процедур их представления, анализа и распознавания, в т.ч. стандартизация разнообразных методов, операций и представлений, используемых в анализе изображений.

## **2. Дескриптивный подход к анализу и пониманию изображений**

На создание ДП оказали существенное влияние следующие основополагающие теории распознавания образов: 1) алгебраический подход к распознаванию образов Ю.И. Журавлева и его алгебра алгоритмов [2]; 2) теория образов У. Гренандера [1]; 3) теория зрения Марра [13]; 4) категорный подход к распознаванию М. Павела [15].

ДП предложен и развивается в качестве концептуальной и логической основы «извлечения информации из изображений». Он объединяет следующие основные инструменты анализа и распознавания изображений: совокупность методов анализа и распознавания изображений, методы приведения изображений к виду, удобному для распознавания (ПВУР [9]), понятийную систему анализа и распознавания изображений [9], классы дескриптивных моделей изображений (ДМИ)[9], язык дескриптивных алгебр изображений (ДАИ)[7], постановки задач анализа и распознавания изображений[8], базовую модель процесса распознавания изображений [5].

Основными направлениями исследований в рамках ДП являются: 1) создание аксиоматики анализа и распознавания изображений; 2) разработка и применение единого языка для описания процессов анализа и распознавания изображений; 3) введение формальных систем на основе некоторых регулярных

структур для определения процессов анализа и распознавания изображений (см., например, [8]).

Математическими основами ДП являются: 1) «алгебраизация» извлечения информации из изображений; 2) специализация Алгебры Журавлева [2] на случай представления исходных данных для распознавания в виде изображений; 3) стандартный язык для описания процедур анализа и распознавания изображений – ДАИ [7]; 4) математическая постановка задачи распознавания изображений [8]; 5) аппарат математической теории анализа изображений и математической теории распознавания образов; 6) модель процесса решения стандартной задачи распознавания изображений.

Основными объектами и средствами ДП являются: 1) изображения; 2) универсальный язык – ДАИ; 3) ДМИ двух типов: а) модели изображений; б) модели процедур решения задач распознавания изображений и их реализации; 4) дескриптивные алгебраические схемы представления изображений [9]; 5) мультимодельные и мультиаспектные представления изображений, которые строятся на основе порождающих дескриптивных деревьев [6].

Базовыми методическими принципами ДП являются: 1) «алгебраизация» анализа изображений; 2) стандартизация представления задач анализа и распознавания изображений; 3) концептуализация и формализация фаз, которые проходит изображение в процессе преобразования при решении задачи распознавания; 4) классификация и спецификация допустимых моделей изображений – ДМИ; 5) ПВУР; 6) использование стандартного алгебраического языка ДАИ для описания моделей изображений и процедур их построения и преобразования; 7) комбинирование алгоритмов в рамках мультиалгоритмических схем; 8) использование мультимодельных и многоаспектных представлений изображений; 9) построение и использование базовой модели процесса решения стандартной задачи распознавания изображений; 10) определение и использование неклассических для математической теории распознавания обра-

зов новых постановок задач анализа и распознавания изображений.

## **Выводы**

Большая часть современных исследований посвящены разработке и исследованию методов математических теорий распознавания образов и анализа изображений. Применение этих методов, как правило, требует построения формальных представлений изображений, допускаемых соответствующими алгоритмами. Исследования различных типов моделей изображений, их погружение в пространство формализации изображений и изучение свойств пространства с привлечением ДАИ являются наиболее перспективными направлениями исследований на сегодняшний момент.

Если эффективное представление изображения – «компактную» модель – построить не удастся, то объектом анализа становится исходное изображение и соответственно возникает задача, для решения которой требуются специальные алгоритмические средства, допускающие пространственные исходные данные. Развитие алгебраической теории анализа изображений идет по пути совместного использования представлений изображений, основанных на комбинаторных структурах локальных окрестностей и признаковых моделях и соответственно совместного использования/комбинирования алгоритмов, допускающих эти представления в качестве исходных данных.

## **Список литературы**

1. *Гренандер У.* Лекции по теории образов, т.1-т.3. – М.: Мир, 1979.
2. *Журавлев Ю. И.* Об алгебраическом подходе к решению задач распознавания и классификации // Проблемы кибернетики, 1978. Вып. 33.С. 5–68.
3. *Barrow H. G., Ambler A. P., Burstall R. M.* Some Techniques for Recognizing Structures in Pictures // *Frontiers of Pattern*

- Recognition (The Proceedings of the International Conference on Frontiers of Pattern Recognition, ed. Satosi Watanabe). – Academic Press, New York, London, 1972. P.1–30.
4. *Evans T. G.* Descriptive Pattern Analysis Techniques: Potentials and Problems // Methodologies of Pattern Recognition (The Proceedings of the International Conference on Methodologies of Pattern Recognition).– Academic Press, New York, London, 1969. -pp. 149-157.
  5. *Gurevich I. B.* Descriptive Technique for Image Description, Representation and Recognition // Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications in the USSR, 1991. Vol. 1. P. 50 – 53.
  6. *Gurevich I., Yashina V.* Generating Descriptive Trees // “Vision, Modeling, and Visualization 2005”, Proceedings, November 16-18, 2005, Erlangen, Germany / G. Greiner, J. Hornegger, H. Niemann, M. Stamminger (Eds.). Infix, 2005. P.367–374
  7. *Gurevich I. B., Yashina V. V.* Operations of Descriptive Image Algebras with One Ring // Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications, 2006. Vol. 16. No. 3. . 298–328.
  8. *Gurevich I. B., Yashina V. V.* Computer-Aided Image Analysis Based on the Concepts of Invariance and Equivalence // Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications, 2006. Vol. 16. No. 4. P. 564–589.
  9. *Gurevich I. B., Yashina V. V.* Image Formalization Space: Formulation of Tasks, Structural Properties, and Elements // Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications, 2011. Vol. 21. No. 2. P.134–139.
  10. *Kaneff S.* Pattern Cognition and the Organization of Information // Frontiers of Pattern Recognition (The Proceedings of the International Conference on Frontiers of Pattern Recognition, ed. Satosi Watanabe). – Academic Press, New York, London, 1972. P. 193–222.

11. *Kirsh R.* Computer Interpretation of English Text and Picture Patterns // IEEE-TEC, Vol. EC-13, No. 4, August, 1964.
12. *Matheron G.* Random Sets and Integral Geometry, Wiley, New York, 1975.
13. *Marr D.* Vision, Freeman, New York, 1982.
14. *Narasimhan R.* Picture Languages // Picture Language Machines (ed. S.Kaneff).— Academic Press, London, New York, 1970. P. 1–30.
15. *Pavel M.* Fundamentals of Pattern Recognition, New York, Marcell, Dekker, Inc., 1989.
16. *Ritter G. X.* Image Algebra. Center for computer vision and visualization, Department of Computer and Information science and Engineering, University of Florida, Gainesville, FL 32611, 2001.
17. *Rosenfeld A.* Picture Languages. Formal Models for Picture Recognition.—Academic Press, New York, San Francisco, London, 1979.
18. *Serra J.* Image Analysis and Mathematical Morphology, London, Academic Press, 1982.
19. *Shaw A.* A Proposed Language for the Formal Description of Pictures // CGS Memo. 28, Stanford University, 1967.
20. *Sternberg S. R.* Grayscale morphology // Computer Vision, Graphics and Image Processing, vol.35, no.3, 1986. P. 333–355.

# Интеграция многокритериальных методов с целью выявления предпочтений

А. И. Бурашникова<sup>68</sup>

**Аннотация:** Статья посвящена выбору трассы трубопровода в Арктике. В работе подробно описаны используемые методы и системы критериев. Результат получен с помощью интеграции предложенных методов.

**Ключевые слова:** ранжирование; интеграция; Арктика; многокритериальность

## Введение

Принятие решений – особый процесс человеческой деятельности, направленный на выбор наилучшего варианта из нескольких возможных [12]. В представленной работе рассмотрено ранжирование альтернатив прокладки трубопровода на полуострове Ямал. Задачи такого рода относятся к классу задач стратегического выбора.

Ранее проведено большое количество исследований применения многокритериальных методов при решении различных задач об освоении Арктики, в частности, разработки месторождений углеводородов, отечественными (Ларичев О.И., Осередько Ю.С. [1]) и зарубежными специалистами (Р.Браун [2, 3, 4, 5]). В вышеуказанных работах использовался только один метод и только одна система критериев. В настоящей статье предполагается интеграция нескольких методов (АНР [8] и ЗАПРОС III [6, 7, 12]) и нескольких систем критериев.

Большинство методов принятия решений базируются на процедурах, не имеющих строгого обоснования. Как правило, подобные процедуры используют опыт и интуицию ЛПР или экспертов. Поэтому возникает необходимость решения конкретной задачи

---

<sup>68</sup>МФТИ, 1993 г.р.; burashnikova@phystech.edu

с помощью различных методов принятия решений (или их комбинаций) и сопоставления полученные результаты. Агрегация нескольких методов и разных систем критериев позволит получить синергетический эффект [11], т.е. более надежное и обоснованное решение.

В результате заболачивания районов крайнего севера, наблюдаемого в последние годы, выработанные ранее технические решения (например, способы прокладки трубопроводов) оказываются неприменимыми в современных условиях. Стоит отметить, что нефтегазовые компании вынуждены использовать методы химической заморозки для укрепления заболачиваемых грунтов, чтобы предотвратить разрывы ниток трубопроводов. Поэтому целесообразно скорректировать старые или даже выработать совершенно новые системы критериев при выборе маршрутов трасс трубопроводов, что существенно важно при использовании современных многокритериальных процедур принятия решений.

## 1. Особенности задач стратегического выбора

Выделим основные характеристики задач стратегического выбора [12]:

- 1) Имеется сравнительно немного вариантов решения проблемы ( $2 \leq |A| \leq 10$ ), из которых следует выбрать один, наилучший.
- 2) Варианты оцениваются по многим критериям ( $|A| < |K|$ ). Среди них могут быть как количественные, так и качественные критерии, при этом последние преобладают.
- 3) Существует большая неопределенность в оценках вариантов по критериям, неустранимая на момент принятия решений.
- 4) Принимаемое решение относится к будущему и его последствия имеют долгосрочный характер.
- 5) Имеется лицо, принимающее решение (ЛПР), несущее основную ответственность за результат принятия решений.
- 6) Задачей ЛПР является выбор наилучшего варианта, соответствующего его целям.

Задача об освоении Арктики удовлетворяет всем этим особенностям. У нас есть небольшое число альтернатив (горный, морской и береговой маршруты), варианты оцениваются по многим критериям (5 и 6), есть ЛППР, задачей которого является выбор наилучшего маршрута трубопровода, присутствует некоторая неопределенность на момент принятия решения и последствия выбора имеют долгосрочный характер.

## 2. Основные характеристики методов, используемых в статье

Подход АНР (предложен Т. Саати) состоит из совокупности этапов [8]:

- 1) Структуризация задачи: цели – критерии – альтернативы
- 2) ЛППР попарно сравнивает элементы каждого уровня, результаты переводятся в числа.
- 3) Подсчет коэффициентов важности для элементов каждого уровня и проверка согласованности суждений ЛППР.
- 4) Подсчет количественного индикатора качества каждой из альтернатив и выбор наилучшей альтернативы (в результате работы метода строится строгий порядок).

Главным достоинством метода АНР является то, что метод АНР может применяться и в тех случаях, когда эксперты не могут дать абсолютные оценки альтернатив по критериям, а пользуются лишь сравнительными измерениями. Главными недостатками является возможное изменение предпочтений между двумя заданными альтернативами при введении новой недоминирующей альтернативы и необоснованный переход к числам при проведении измерений [9, 10]. Что касается метода ЗАПРОС III (семейство методов ЗАПРОС – О. И. Ларичев, 1978 г.), то он состоит из совокупности следующих этапов:

- 1) Считаем, что нам заданы критерии оценки альтернатив с вербальными оценками на шкалах

- 2) Опираясь на п. 1., строится правило упорядочения многокритериальных альтернатив с использованием парных сравнений
- 3) На основе парных сравнений строится единая порядковая шкала (ЕШИК)
- 4) Ранжирование заданных альтернатив с использованием полученного правила (в результате работы метода строится квазипорядок)

Главным достоинством метода ЗАПРОС III является использование психологически корректных процедур выявления предпочтений, что позволяет учесть возможности человека при переработке информации, в том числе различные ошибки (противоречия). Но по сравнению с АНР, результаты в ЗАПРОС III носят приближенный характер (альтернативы получают ранги, а не количественные значения ценности) [6, 7, 12].

Указанные недостатки предполагают одновременное использование этих методов при решении задачи, с целью их устранения в процессе интеграции результатов.

### **3. Постановка задачи об освоении Арктики**

Проблема выбора трассы газопровода на полуострове Ямал рассматривается как одна из распространенных задач принятия решений, а именно, ранжирования многокритериальных альтернатив. Предлагается три возможных варианта: вдоль берега Ямальского полуострова (береговой вариант), пересечение Байдарацкой губы и залива Карского моря (морской вариант) и вдоль мыса Каменный (горный вариант). Береговой вариант является более определенным по условиям строительства, но наносит большой ущерб окружающей среде. Морской вариант значительно короче, чем береговой и горный, но связан с существенной неопределенностью и отсутствием опыта строительства. Горный вариант самый протяженный, но срок строительства подобного маршрута фиксирован. Также, этот маршрут оказывает гораздо меньшее влияние на окружающую среду, чем береговой маршрут.



Рис. 1. Морской, береговой и горный маршруты

#### 4. Введенные критерии в методе ЗАПРОС III

##### 1) Затраты

Этот критерий позволяет нам выбрать трассу, которая потребует минимальных затрат, что однако не гарантирует нам выбор оптимальной трассы, так как при этом не учитывается ряд других факторов.

##### 2) Срок строительства

Этот фактор может оказаться решающим, если главной задачей является быстрый ввод трубопровода в эксплуатацию.

##### 3) Влияние на окружающую среду

Сооружение трубопроводов оказывает большое влияние на окружающую среду. Например, в нашем случае, при прокладке подводных трубопроводов с недостаточным углублением в грунт, возникают подземные деформации, что отрицательно сказывается на безаварийном состоянии трубопровода.

##### 4) Удобство эксплуатации

Чтобы обеспечить безопасную работу трубопровода, необходимо иметь доступ ко всем его участкам. Возможность доступа определяется природными условиями в районе трассы.

#### 5) Влияние на население

Здесь необходимо учитывать аспект, связанный с необходимостью сноса домов для прокладки трубопровода, а также другие отрицательные эффекты, напрямую касающиеся населения.

Согласно указанным критериям, было проведено ранжирование альтернатив. Всего получилось 119 абсолютных рангов.

**Таблица 1.** Итоговый результат по альтернативам (ЗАПРОС III)

№	Название	Ранг
1	Морской	45
2	Береговой	76
3	Горный	82

Исходя из этого, делаем вывод, что морской вариант маршрута более предпочтителен, чем береговой и горный.

## 5. Введенные критерии в методе АНР

В методе АНР используется усовершенствованная система критериев, в нее добавлена оценка по надежности эксплуатации. Таким образом, если в предыдущем примере для нас большее значение имела скорость строительства трубопровода, то здесь, мы акцентируем свое внимание на впоследствии долговременной и надежной работе трубопровода.

## 6. Результат ранжирования альтернатив методом АНР

В результате ранжирования, каждой альтернативе были сопоставлены веса, указанные в таблице.

**Таблица 2.** Итоговый результат по альтернативам(АНР)

№	Название	Вес
1	Морской	0.505
2	Береговой	0.304
3	Горный	0.191

## Выводы

В представленной статье проведена интеграция двух разных методов - АНР и ЗАПРОС III, а также двух разных систем критериев с целью ранжирования альтернатив на примере задачи о прокладке трубопровода на полуострове Ямал. В результате объединения двух подходов был выбран наиболее предпочтительный маршрут. В последующих работах предполагается расширить область используемых методов и систем критериев. Для будущей агрегации результатов планируется использовать теорию решеток (Биркгофф Г.) и упорядочение по предпочтениям (Кемени-Снелл).

## Благодарности

*Хочу выразить глубокую благодарность Ройзензону Г.В. за помощь и постоянный интерес к работе, а также за ценные советы и замечания.*

## Список литературы

1. *Осередько Ю. С., Ларичев О. И., Мечитов А. И.* Исследование процесса выбора трассы магистрального трубопровода // Проблемы и процедуры принятия решений при многих критериях. – М.: ВНИИСИ, 1982. С. 61–72.
2. *Ларичев О. И., Браун Р. В.* Количественный и вербальный анализ решений: сравнительное исследование возможностей и ограничений // Экономика и математические методы, 1998. №4. С. 97–107.

3. *Flanders N., Brown R., Andreeva Y., Larichev O.* Justifying Public Decisions in Arctic Oil and Gas Development: US and Russian Approaches // *Arctic*, 1998. Vol.51. No.3. Pp. 262–279.
4. *Larichev O., Brown R.*, Numerical and Verbal Decision Analysis used for the problems of resources allocation in Arctic // *Journal of Multi-Criteria Decision Analysis*, 2000. Vol.9. No.6. Pp. 263–274.
5. *Larichev O. I., Brown R. W.* Comparing Numerical and Verbal Decision Analysis Using an Arctic Resource Management Problem // *Multiple Criteria Decision Making in the New Millennium*, 2001. Vol.507 of *Lecture Notes in Economics and Mathematical Systems*. Pp. 79–92.
6. *Димитриади Г. Г., Ларичев О. И.* Система поддержки принятия решений и метод ЗАПРОС-III: ранжирование многокритериальных альтернатив с вербальными оценками качества // *Автоматика и телемеханика.*, 2005. No.8. С. 146–160.
7. *Ларичев О. И., Зуев Ю. А., Гнеденко Л. С.* Метод ЗАПРОС (Замкнутые Процедуры у Опорных Ситуаций) анализа вариантов сложных решений // *Многокритериальный выбор при решении слабоструктуризованных задач* . – М.: ВНИИСИ, 1978. С. 83–97.
8. *Saaty T.* Принятие решений. Метод анализа иерархий. – М.: Радио и связь, 1993. – 278 с.
9. *Belton V., Gear A. E.* On a short-coming of Saaty’s method of analytical hierarchies // *Omega*, 1983. Vol.11. No.3. Pp. 228–230.
10. *Triantaphyllou E.* Two new cases of rank reversals when the AHP and some of its additive variants are used that do not occur with the multiplicative AHP // *Journal of Multi-criteria Decision Analysis*, 2001. Vol.10. No.1. Pp. 11–25.
11. *Ройзензон Г. В.* Синергетический эффект в принятии решений // *Системные исследования. Методологические проблемы. Ежегодник*. – М.: УРСС, 2012. С. 248–272.
12. *Ларичев О. И.* Вербальный анализ решений. – М.: Наука, 2006. – 181 с.

# Оценки скорости сходимости в центральных предельных теоремах для случайных сумм при ослабленных моментных условиях\*

А. В. Дорофеева<sup>69</sup>, В. Ю. Королев<sup>70</sup>

**Аннотация:** Получены оценки точности нормальной аппроксимации для пуассон-биномиальных, биномиальных и пуассоновских случайных сумм при ослабленных моментных условиях, в частности, при отсутствии моментов порядков, больших второго. Также выписаны оценки точности приближения распределений смешанных пуассоновских случайных сумм соответствующими предельными законами. В частности, получены оценки точности аппроксимации геометрических, отрицательных биномиальных и зихелевых случайных сумм распределениями Лапласа, дисперсионным гамма и Стьюдента соответственно. Оценки всех абсолютных констант выписаны в явном виде.

**Ключевые слова:** центральная предельная теорема; нормальное распределение; оценка скорости сходимости; равномерная метрика; пуассон-биномиальное распределение; пуассон-биномиальная случайная сумма; биномиальная случайная сумма; пуассоновская случайная сумма; смешанная пуассоновская случайная сумма; геометрическая случайная сумма; отрицательная биномиальная случайная сумма; смешанная пуассоновская случайная сумма; распределение Зихеля; распределение Лапласа; дисперсионное гамма-распределение; распределение Стьюдента; абсолютная константа

## Введение

На сегодняшний день большое количество моделей, применяемых в различных областях человеческой деятельности, базируется на использовании случайных сумм. Так, например,

---

\*Работа выполнена при поддержке Российского фонда Фундаментальных исследований, грант №15-07-02984

<sup>69</sup>ВМК МГУ им. М. В. Ломоносова; 1991 г.р.; alex.dorofeyeva@gmail.com

<sup>70</sup>ВМК МГУ им. М. В. Ломоносова, ФИЦ ИУ РАН (ИПИ РАН); 1954 г.р.; д.ф.-м.н.; vkorolev@cs.msu.ru

пуассон-биномиальные, биномиальные и смешанные пуассоновские (в первую очередь геометрические) случайные суммы широко применяются при описании остановленных случайных блужданий в финансовой математике, страховании, теории надежности.

Наличие большого количества практических приложений у упомянутых выше случайных сумм приводит к необходимости получения оценок точности их аппроксимаций. Более того, принято считать, что распределение элементарных скачков случайных блужданий может иметь очень тяжелые хвосты. Задачи, рассматриваемые в настоящей работе, соответствуют ситуациям, когда хвосты могут быть сколь угодно тяжелыми при условии адекватности нормальной аппроксимации. Кроме того, оценки, полученные в данной работе, дают ответ на вопрос, насколько тяжелыми могут быть эти хвосты для того, чтобы нормальная аппроксимация (масштабно-смешанная нормальная аппроксимация) была адекватна.

## 1. Точность нормальной аппроксимации для распределений пуассон-биномиальных случайных сумм

Пусть  $X_1, X_2, \dots$  – одинаково распределенные случайные величины с  $EX_i = 0$  и  $0 < EX_i^2 \equiv \sigma^2 < \infty$ . Пусть  $p_j \in (0, 1]$  – произвольные числа,  $j = 1, 2, \dots$ . Для  $n \in \mathbb{N}$  обозначим  $\theta_n = p_1 + \dots + p_n$ ,  $\mathbf{p}_n = (p_1, \dots, p_n)$ . Распределение случайной величины

$$N_{n, \mathbf{p}_n} = \xi_1 + \dots + \xi_n,$$

где  $\xi_1, \dots, \xi_n$  – независимые случайные величины такие, что

$$\xi_j = \begin{cases} 1 & \text{с вероятностью } p_j, \\ 0 & \text{с вероятностью } 1 - p_j, \end{cases}, \quad j = 1, \dots, n,$$

называется пуассон-биномиальным распределением с параметрами  $n$  и  $\mathbf{p}_n$ . Пусть при каждом  $n \in \mathbb{N}$  случайные величины  $N_{n, \mathbf{p}_n}, X_1, X_2, \dots$  независимы в совокупности. Основным объектом, рассматриваемым в данном разделе, являются *пуассон-биномиальные случайные суммы* вида:  $S_{N_{n, \mathbf{p}_n}} = X_1 + \dots + X_{N_{n, \mathbf{p}_n}}$ . При этом, если  $N_{n, \mathbf{p}_n} = 0$ , то полагаем  $S_{N_{n, \mathbf{p}_n}} = 0$ .

Заметим, что  $DS_{N_n, p_n} = \theta_n \sigma^2$ , и введем следующее обозначение:  $\Delta_{n, p_n} = \sup_x |\mathbb{P}(S_{N_n, p_n} < x\sigma\sqrt{\theta_n}) - \Phi(x)|$ . Тогда имеет место следующая

ТЕОРЕМА 1. Для любых  $n \in \mathbb{N}$  и  $p_j \in (0, 1]$ ,  $j \in \mathbb{N}$ ,

$$\Delta_{n, p_n} \leq \frac{1.8627}{\sigma^2} \mathbb{E}X_1^2 \min \left\{ 1, \frac{|X_1|}{\sigma\sqrt{\theta_n}} \right\}.$$

Введем класс  $\mathcal{G}$  вещественных функций  $g(x)$  аргумента  $x \in \mathbb{R}$  таких, что 1)  $g(x)$  четна; 2)  $g(x)$  неотрицательна при всех  $x$ , а также  $g(x) > 0$  при всех  $x > 0$ ; 3)  $g(x)$  и  $x/g(x)$  не убывают при  $x > 0$ .

Тогда справедлива следующая

ТЕОРЕМА 2. В условиях теоремы 1, какой бы ни была функция  $g \in \mathcal{G}$  такая, что  $\mathbb{E}X_1^2 g(X_1) < \infty$ , справедливо неравенство

$$\Delta_{n, p_n} \leq 1.8627 \frac{\mathbb{E}X_1^2 g(X_1)}{\sigma^2 g(\sigma\sqrt{\theta_n})}.$$

В частности, если  $p_1 = p_2 = \dots = p$ , то пуассон-биномиальное распределение с параметрами  $n \in \mathbb{N}$  и  $p_n$  становится классическим биномиальным распределением с параметрами  $n$  и  $p$ :

$$N_{n, p_n} \stackrel{d}{=} N_{n, p}, \quad \mathbb{P}(N_{n, p} = k) = C_n^k p^k (1-p)^{n-k}, \quad k = 0, \dots, n,$$

где символ  $\stackrel{d}{=}$  обозначает совпадение распределений.

При этом легко видеть, что  $\theta_n = np$ , так что  $DS_{N_n, p} = np\sigma^2$ . Обозначим  $\Delta_{n, p} = \sup_x |\mathbb{P}(S_{N_n, p} < x\sigma\sqrt{np}) - \Phi(x)|$ . Важными следствиями являются:

СЛЕДСТВИЕ 1. Для любых  $n \in \mathbb{N}$  и  $p \in (0, 1]$

$$\Delta_{n, p} \leq \frac{1.8546}{\sigma^2} \mathbb{E}X_1^2 \min \left\{ 1, \frac{|X_1|}{\sigma\sqrt{np}} \right\}.$$

СЛЕДСТВИЕ 2. В условиях теоремы 1, какой бы ни была функция  $g \in \mathcal{G}$  такая, что  $\mathbb{E}X_1^2 g(X_1) < \infty$ , при любых  $n \in \mathbb{N}$  and  $p \in (0, 1]$  справедливо следующее неравенство

$$\Delta_{n, p_n} \leq 1.8546 \frac{\mathbb{E}X_1^2 g(X_1)}{\sigma^2 g(\sigma\sqrt{np})}.$$

## 2. Точность нормальной аппроксимации для распределений пуассоновских случайных сумм

В дополнение к обозначениям, введенным выше, пусть  $\lambda > 0$  и  $N_\lambda$  – случайная величина, имеющая распределение Пуассона с параметром  $\lambda$ :  $P(N_\lambda = k) = e^{-\lambda} \frac{\lambda^k}{k!}$ ,  $k \in \mathbb{N} \cup \{0\}$ .

Предположим, что при любом  $\lambda > 0$  случайные величины  $N_\lambda, X_1, X_2, \dots$  независимы в совокупности. Рассмотрим пуассоновскую случайную сумму вида:  $S_{N_\lambda} = X_1 + \dots + X_{N_\lambda}$ . При этом, если  $N_\lambda = 0$ , то полагаем  $S_{N_\lambda} = 0$ . Несложно видеть, что  $ES_\lambda = 0$  и  $DS_\lambda = \lambda\sigma^2$ . Обозначим  $\Delta_\lambda = \sup_x |P(S_\lambda < x\sigma\sqrt{\lambda}) - \Phi(x)|$ . Имеют место следующие утверждения:

ТЕОРЕМА 3. Для любого  $\lambda > 0$

$$\Delta_\lambda \leq \frac{1.8546}{\sigma^2} EX_1^2 \min \left\{ 1, \frac{|X_1|}{\sigma\sqrt{\lambda}} \right\}.$$

ТЕОРЕМА 4. Какой бы ни была функция  $g \in \mathcal{G}$  такая, что  $EX_1^2 g(X_1) < \infty$ , справедливо неравенство

$$\Delta_\lambda \leq 1.8546 \frac{EX_1^2 g(X_1)}{\sigma^2 g(\sigma\sqrt{\lambda})}. \quad (1)$$

ЗАМЕЧАНИЕ 1. Очевидно, что  $g(x) \equiv |x| \in \mathcal{G}$ . Для такой функции  $g$  неравенство (1) принимает вид классического неравенства Берри – Эссеена для пуассоновских сумм, наилучшая на сегодняшний день верхняя оценка абсолютной константы для которого приведена в работе [1]:

$$\Delta_\lambda \leq 0.3031 \frac{E|X_1|^3}{\sigma^3\sqrt{\lambda}}.$$

### 3. Оценки скорости сходимости смешанных пуассоновских случайных сумм к соответствующим распределениям

Введем «бесконечно большой» параметр  $n \in \mathbb{N}$  и рассмотрим случайные величины  $N_n^*$  такие, что для каждого  $n \in \mathbb{N}$

$$P(N_n^* = k) = \int_0^\infty e^{-\lambda} \frac{\lambda^k}{k!} dP(\Lambda_n < \lambda), \quad k \in \mathbb{N} \cup \{0\}, \quad (2)$$

где  $\Lambda_n$  – некоторая положительная случайная величина. Для простоты  $n$  может быть рассмотрено как параметр масштаба распределения  $\Lambda_n$ , т. е.  $\Lambda_n = n\Lambda$ , где  $\Lambda$  – некоторая положительная «стандартная» случайная величина, в том смысле, что  $E\Lambda = 1$  (если последнее существует).

Пусть для любого  $n \in \mathbb{N}$  случайные величины  $N_n^*$  и  $X_1, X_2, \dots$  независимы в совокупности. Пусть  $S_{N_n^*} = X_1 + \dots + X_{N_n^*}$ . Если  $N_n^* = 0$ , то  $S_{N_n^*} = 0$ . Согласно (2) легко видеть, что если  $E\Lambda_n < \infty$ , то  $EN_n^* = E\Lambda_n$  и  $DS_n = \sigma^2 E\Lambda_n$ .

Пусть  $N_\lambda$  – пуассоновская случайная величина с параметром  $\lambda$ , не зависящая от  $X_1, X_2, \dots$ . Обозначим

$$\Delta_n^* = \sup_x \left| P(S_{N_n^*} < x\sigma\sqrt{E\Lambda_n}) - \int_0^\infty \Phi\left(\frac{x}{\sqrt{\lambda}}\right) dP(\Lambda_n < \lambda E\Lambda_n) \right|.$$

Имеет место следующая

ТЕОРЕМА 5. Если  $E\Lambda_n < \infty$ , то

$$\begin{aligned} \Delta_n^* &\leq \frac{1.8546}{\sigma^2} EX_1^2 \min \left\{ 1, \frac{|X_1|}{\sigma\sqrt{\Lambda_n}} \right\} = \\ &= \frac{1.8546}{\sigma^2} \left[ EX_1^2 \mathbb{I}(|X_1| \geq \sigma\sqrt{\Lambda_n}) + E \frac{|X_1|^3}{\sigma\sqrt{\Lambda_n}} \mathbb{I}(|X_1| < \sigma\sqrt{\Lambda_n}) \right], \end{aligned}$$

где случайные величины  $X_1$  и  $\Lambda_n$  предполагаются независимыми.

Задавая распределение  $N_n^*$ , из упомянутой выше теоремы можно получить важные следствия.

Так, например, в работе рассмотрен случай, когда  $N_n^*$  имеет геометрическое распределение с параметром  $p = \frac{1}{1+n}$ ,  $n \in \mathbb{N}$ :

$$P(N_n^* = k) = \frac{1}{n+1} \left( \frac{n}{n+1} \right)^k, \quad k \in \mathbb{N} \cup \{0\},$$

и получены оценки скорости сходимости геометрических случайных сумм к распределению Лапласа.

Также выписан результат для ситуации, когда  $N_n^*$  имеет отрицательное биномиальное распределение с параметрами  $r > 0$  и  $\frac{1}{n+1}$ :

$$\begin{aligned} P(N_n^* = k) &= \frac{1}{n^r \Gamma(r)} \int_0^\infty e^{-\lambda} \frac{\lambda^k}{k!} \lambda^{r-1} e^{-\lambda/n} d\lambda = \\ &= \frac{\Gamma(r+k)}{\Gamma(r) k!} \left( \frac{1}{1+n} \right)^r \left( \frac{n}{1+n} \right)^k, \quad k \in \mathbb{N} \cup \{0\}, \end{aligned}$$

и получены оценки скорости сходимости отрицательных биномиальных случайных сумм к дисперсионному гамма распределению.

Более того, рассмотрен случай, когда в представлении (2)  $\Lambda_n$  – случайная величина, имеющая обратное гамма-распределение с параметрами  $\frac{r}{2}$  и  $\frac{n}{2}$  с плотностью:

$$p(\lambda) = \frac{n^{r/2} \lambda^{-r/2-1}}{2^{r/2} \Gamma(\frac{r}{2})} \exp \left\{ -\frac{n}{2\lambda} \right\}, \quad \lambda > 0, \quad r > 1.$$

Тогда случайная величина  $N_n^*$  имеет так называемое пуассон-обратное гамма-распределение:

$$P(N_n^* = k) = \frac{n^{r/2}}{2^{r/2} \Gamma(\frac{r}{2})} \int_0^\infty e^{-\lambda} \frac{\lambda^k}{k!} \lambda^{-r/2-1} \exp \left\{ -\frac{n}{2\lambda} \right\} d\lambda, \quad k \in \mathbb{N} \cup \{0\},$$

которое является частным случаем так называемого распределения Зихеля (см., например, [2, 3]). В этом случае представлены оценки скорости сходимости распределений зихелевых случайных сумм к распределению Стьюдента.

## Выводы

В работе получены оценки точности нормальной аппроксимации для пуассон-биномиальных, биномиальных и пуассоновских

случайных сумм при ослабленных моментных условиях, в частности, при отсутствии моментов порядков, больших второго. Также в качестве частных случаев выписаны оценки точности аппроксимаций: 1) геометрических случайных сумм распределением Лапласа; 2) отрицательных биномиальных случайных сумм дисперсионным гамма распределением; 3) зихелевых случайных сумм распределением Стьюдента.

## Список литературы

1. *Шевцова И. Г.* Об абсолютных константах в неравенствах типа Берри-Эссеена // Доклады Академии наук, 2014. Т. 456. Вып. 6. С. 650–654.
2. *Sichel H. S.* On a family of discrete distributions particular suited to represent long tailed frequency data // Laubscher N. F. (Ed.). Proceedings of the 3rd Symposium on Mathematical Statistics. – Pretoria: CSIR, 1971. P. 51–97.
3. *Willmot G. E.* On recursive evaluation of mixed Poisson probabilities and related quantities // Scandinavian Actuarial Journal, 1993, No. 2. P. 114–133.
4. *Sunklodas J. K.* On the normal approximation of a binomial random sum // Lithuanian Mathematical Journal, 2014. Vol. 54. No. 3. DOI:10.1007/s10986-014-9248-6.
5. *Королев В. Ю., Попов С. В.* Уточнение оценок скорости сходимости в центральной предельной теореме при отсутствии моментов порядков, больших второго // Теория вероятностей и ее применения, 2011. Т. 56. Вып. 4. С. 797–805.
6. *Королев В. Ю., Дорофеева А. В.* Оценки точности нормального приближения для распределений случайных сумм при ослабленных моментных условиях // В сб. «Статистические методы оценивания и проверки гипотез». Вып. 26. – Пермь: Изд-во Пермского государственного национального исследовательского университета, 2015.

# Оценка параметров линейной регрессии методом максимальной энтропии

Ю. А. Дубнов<sup>71</sup>

**Аннотация:** Построение регрессионных моделей является одной из фундаментальных задач статистики и эконометрики. Работа посвящена исследованию метода максимальной энтропии для оценки параметров регрессии. В работе приведено описание метода, рассмотрены некоторые его модификации и проведено качественное сравнение с наиболее распространенными альтернативами на примере нормальных и смещенных возмущений.

**Ключевые слова:** метод максимальной энтропии; линейная регрессия; байесовский подход

## Введение

Принцип максимизации энтропии впервые был сформулирован Эдвардом Т. Джейнсом в 1957 году в качестве метода приближенного решения переопределенных систем уравнений. И только в 1996 году А. Голан и др. предложили использовать этот принцип для построения регрессионных моделей в эконометрике, представив метод максимальной энтропии (GME, Generalized Maximum Entropy) [1].

Как и другие известные методы оценивания, такие как, метод максимального правдоподобия (ММП) и байесовский подход, метод максимальной энтропии (ММЭ) основан на предположении о стохастическом характере модели. Так, параметрами модели полагаются случайные величины, а их вероятностные характеристики восстанавливаются в процессе оценивания. В частности, в ММЭ восстанавливаются энтропийно-оптимальные плотности вероятности параметров модели, реализующие максимум информационной энтропии Больцмана [2], которые далее используются для построения точечных и интервальных оценок параметров.

---

<sup>71</sup>ФИЦ ИУ РАН (ИСА РАН); МФТИ; 1990 г.р.; yury.dubnov@phystech.edu

В работах [3], [4] рассмотрен классический метод максимальной энтропии (ГМЕ) в дискретном исполнении, авторы приводят результаты экспериментов, демонстрирующие эффективность ММЭ для разного типа ошибок по сравнению с методом наименьших квадратов. В этих экспериментах не использовалась априорная информация об ошибках, что эквивалентно равномерному априорному распределению. Далее рассмотрена непрерывная реализация ММЭ с априорными распределениями ошибок, заданными по нормальному закону и закону «хи-квадрат».

## 1. Описание метода максимальной энтропии

Линейная регрессия задается системой линейных уравнений:

$$\mathbf{y} = \mathbf{X}\bar{\beta} + \bar{\varepsilon}, \quad (1)$$

где  $\mathbf{y}$  - заданный  $T \times 1$  вектор ответов для точек, заданных  $T \times K$  матрицей  $\mathbf{X}$  значений факторов,  $\bar{\beta}$  -  $K \times 1$  вектор искомых параметров, а  $\bar{\varepsilon}$  -  $T \times 1$  вектор ошибок измерений.

Пусть  $p_0^{(k)}$  и  $q_0^{(t)}$  - априорные плотности для параметров  $\beta_k$  и шумов  $\varepsilon_t$ , заданные на интервалах  $B^{(k)}$  и  $E^{(t)}$  соответственно. Задача максимизации энтропии в этом случае записывается с использованием понятия расстояния Кульбака:

$$H(P, Q) = - \int_{\mathbf{B}} P(\bar{\beta}) \ln \frac{P(\bar{\beta})}{P_0(\bar{\beta})} d\bar{\beta} - \int_{\mathbf{E}} Q(\bar{\varepsilon}) \ln \frac{Q(\bar{\varepsilon})}{Q_0(\bar{\varepsilon})} d\bar{\varepsilon} \Rightarrow \max_{P, Q}, \quad (2)$$

при ограничениях

$$\int_{\mathbf{B}} \bar{X}'_k \bar{\beta} P(\bar{\beta}) d\bar{\beta} + \int_{\mathbf{E}} \varepsilon_t Q(\bar{\varepsilon}) d\bar{\varepsilon} = y_t, \quad t = \overline{1, T},$$

$$\int_{\mathbf{B}} P(\bar{\beta}) d\bar{\beta} = 1, \quad \int_{\mathbf{E}} Q(\bar{\varepsilon}) d\bar{\varepsilon} = 1,$$

где  $\bar{X}'_k$  -  $k$ -ый столбец матрицы  $\mathbf{X}$  и, с учетом независимости случайных величин совместные плотности  $P(\bar{\beta})$  и  $Q(\bar{\varepsilon})$  распадаются на

произведения компонент  $p^{(k)}(\beta_k)$  и  $q^{(t)}(\varepsilon_t)$  соответственно. Решение экстремальной задачи (2) имеет вид [5]:

$$p^{(k)}(\beta_k) = \frac{p_0^{(k)} e^{-\beta_k \sum_t \lambda_t X_{tk}}}{\int_{B_k} p_0^{(k)} e^{-\beta_k \sum_t \lambda_t X_{tk}} d\beta_k}, \quad q^{(t)}(\varepsilon_t) = \frac{q_0^{(t)} e^{-\lambda_t \varepsilon_t}}{\int_{E_t} q_0^{(t)} e^{-\lambda_t \varepsilon_t} d\varepsilon_t}. \quad (3)$$

Множители Лагранжа  $\bar{\lambda} = \{\lambda_t, t = \overline{1, T}\}$  являются решением системы балансовых уравнений – первое ограничение в (2).

Таким образом, аналогично байесовскому подходу, априорные вероятности  $p_0^{(k)}$  и  $q_0^{(t)}$  уточняются в соответствии с наблюдениями. Но если в байесовском методе критерием оптимальности служит теорема Байеса, то здесь – максимум энтропии или минимум относительной энтропии (relative entropy).

Наконец, точечные оценки параметров модели (1) вычисляются по формуле:

$$\hat{\beta}_k = \mathbf{E}^{\mathbf{P}}\{\beta_k\} = \int_{B_k} \beta_k p^{(k)}(\beta_k) d\beta_k. \quad (4)$$

## 2. Сравнение с другими методами

Сравнение проводится на примере парной регрессии  $y_t = \alpha + \beta x_t + \varepsilon_t$ ,  $t = \overline{1, T}$ . Исследуется точность оценок параметра  $\hat{\beta}$  в зависимости от объема входных данных  $T$  с нормальными и смещенными ошибками.

### 2.1. Нормально распределенные ошибки

Ошибки, традиционно задаваемые при построении регрессионных моделей [6]:

$$\varepsilon_t \propto N(0, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\varepsilon_t^2}{2\sigma^2}\right), \quad t = \overline{1, T}. \quad (5)$$

Известно, что в этом случае точечные оценки параметров как по методу максимального правдоподобия [6], так и по байесовскому методу [7], являются несмещенными и совпадают с оценками метода наименьших квадратов.

$$\hat{\beta}_{ls} = \frac{\sum_t (y - \bar{y})(x - \bar{x})}{\sum_t (x - \bar{x})^2} = \frac{\bar{x}\bar{y} - \bar{x}\bar{y}}{\bar{x}^2 - \bar{x}^2}, \quad \hat{\alpha}_{ls} = \bar{y} - \hat{\beta}_{ls}\bar{x} \quad (6)$$

При этом априорные распределения параметров полагаются равномерными, т.е.  $p_0(\alpha) = Const$ ,  $p_0(\beta) = Const$ . Тогда выражения в (3), задающие энтропийно-оптимальные распределения, примут вид:

$$p(\alpha) = \frac{e^{-\alpha \sum_t \lambda_t}}{\int_A e^{-\alpha \sum_t \lambda_t} d\alpha}, \quad p(\beta) = \frac{e^{-\beta \sum_t \lambda_t x_t}}{\int_B e^{-\beta \sum_t \lambda_t x_t} d\beta},$$

$$q^{(t)}(\varepsilon_t) = \frac{e^{-\lambda_t \varepsilon_t - \frac{\varepsilon_t^2}{2\sigma^2}}}{\int_{E_t} e^{-\lambda_t \varepsilon_t - \frac{\varepsilon_t^2}{2\sigma^2}} d\varepsilon_t}, \quad (7)$$

где  $A, B$  - выбранные интервалы значений параметров  $\alpha, \beta$ , интервал шумов  $E_t = (-\infty, \infty)$ .

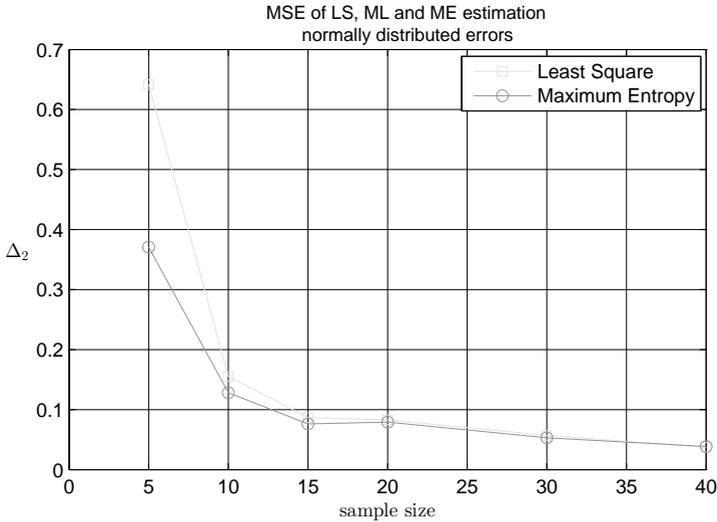
Для сравнения оценок был проведен эксперимент, включающий генерацию массива данных с выбранными параметрами, разбиение начальных данных ошибками с выбранным законом распределения, вычисление и усреднение оценок. Количественное сравнение проводится по метрике  $L_2$ , характеризующей среднеквадратическое отклонение (MSE, mean square error) оценок  $\hat{\beta}$  от своего истинного значения  $\beta^*$ :

$$\Delta_2 = \frac{1}{N} \sum_{i=1}^N (\hat{\beta} - \beta^*)^2.$$

На Рис. 1 приведены результаты экспериментов, демонстрирующие преимущество ММЭ в условиях малого объема входных данных. В то время как монотонное убывание MSE с ростом объема выборки говорит об асимптотической состоятельности оценок.

## 2.2. Ошибки «хи-квадрат»

Пусть ошибки измерений задаются по закону  $\chi^2$  с числом степеней свободы  $f$ :



**Рис. 1.** Среднеквадратическое отклонение оценок

$$\varepsilon_t \propto \chi^2(f) = \frac{(\frac{1}{2})^{\frac{f}{2}}}{\Gamma(\frac{f}{2})} \varepsilon_t^{\frac{f}{2}-1} \exp(-\frac{\varepsilon_t}{2}),, \quad t = \overline{1, T}, \quad (8)$$

где  $\Gamma(\frac{f}{2})$  - гамма-функция.

Оценки по методу наименьших квадратов не «привязаны» к закону распределения ошибок и по-прежнему вычисляются по формуле (6). В то время, как в ММП и байесовском подходе используется функция правдоподобия, принимающая с учетом (8) вид:

$$L(\mathbf{y}, \mathbf{x} | \alpha, \beta, f) = \prod_{t=1}^T \frac{(\frac{1}{2})^{\frac{f}{2}}}{\Gamma(\frac{f}{2})} (y_t - \alpha - \beta x_t)^{\frac{f}{2}-1} \exp\left(-\frac{(y_t - \alpha - \beta x_t)}{2}\right) \quad (9)$$

Максимизация  $L$  определяет оценки параметров  $\theta = \{\alpha, \beta, f\}$  по методу максимального правдоподобия,  $\hat{\theta}_{ml} = \arg \max_{\theta} \ln L(\mathbf{y}, \mathbf{x} | \theta)$  В байесовском методе функция правдоподобия используется при восстановлении апостериорных распределений по формуле Байеса [7]:

$$p(\theta|\mathbf{y}, \mathbf{x}) = \frac{L(\mathbf{y}, \mathbf{x}|\theta)p_0(\theta)}{p(\mathbf{y}, \mathbf{x})}, \quad p(\mathbf{y}, \mathbf{x}) = \int L(\mathbf{y}, \mathbf{x}|\theta)p_0(\theta)d\theta$$

В общем случае апостериорное распределение отличается от функции правдоподобия, но в предположении о расплывчатой априорной информации, совпадает с ней с точностью до нормировочной константы, то есть максимум апостериорного распределения (МАР), определяющий байесовские оценки, совпадает с максимумом функции правдоподобия:  $\hat{\theta}_{bayes} = \arg \max_{\theta} p(\theta|\mathbf{y}, \mathbf{x}) = \hat{\theta}_{ml}$ .

Для вычисления ММЭ-оценок, подставим (8) в выражение (3):

$$p(\alpha) = \frac{e^{-\alpha \sum_t \lambda_t}}{\int_A e^{-\alpha \sum_t \lambda_t} d\alpha}, \quad p(\beta) = \frac{e^{-\beta \sum_t \lambda_t x_t}}{\int_B e^{-\beta \sum_t \lambda_t x_t} d\beta},$$

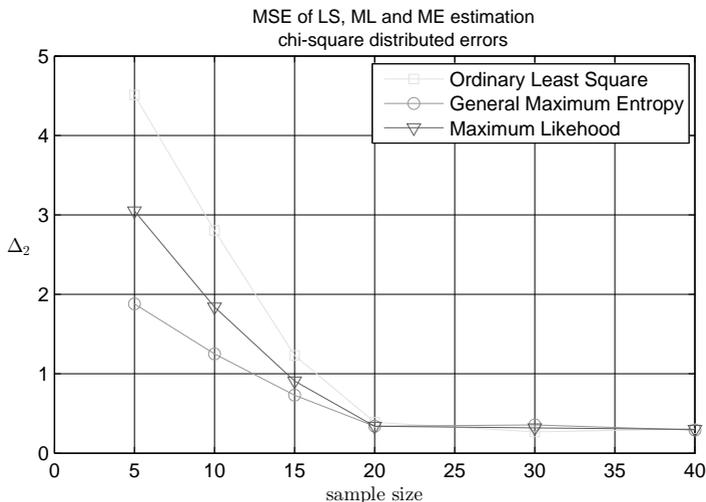
$$q^{(t)}(\varepsilon_t) = \frac{\varepsilon_t^{\frac{f}{2}-1} e^{-\lambda_t \varepsilon_t - \frac{\varepsilon_t}{2}}}{\int_{E_t} \varepsilon_t^{\frac{f}{2}-1} e^{-\lambda_t \varepsilon_t - \frac{\varepsilon_t}{2}} d\varepsilon_t}, \quad (10)$$

В случае распределения «хи-квадрат» интервал шума  $E_t = [0, \infty)$ . Аналогично случаю нормально распределенных ошибок, здесь априорная плотность, задаваемая законом «хи-квадрат», преобразуется в энтропийно-оптимальную плотность шума.

Результаты экспериментов представлены на Рис. 2. В этом случае оценки всех методов оказываются смещенными, в то время как именно метод максимальной энтропии показывает наиболее низкие значения дисперсии ошибок в условиях малого объема данных.

## Выводы

В работе описан метод максимальной энтропии для оценки параметров линейной регрессионной модели, рассмотрен механизм построения энтропийно-оптимальных плотностей и исследована зависимость ММЭ-оценок от объема начальных данных. В ходе численных экспериментов показана эффективность метода максимальной энтропии при ограниченном объеме данных.



**Рис. 2.** Среднеквадратическое отклонение оценок.

## Список литературы

1. *Amos Golan, George G. Judge, Douglas Miller.* Maximum Entropy Econometrics: Robust Estimation with Limited Data. – John Wiley and Sons Ltd. Chichester, U.K., 1996. – 324 p.
2. *Cover T. M., Thomas J. A.* Elements of information theory. – John Wiley and Sons Ltd, New York, 1991. – 561 p.
3. *H. Ozan Eryugur.* Generalized maximum entropy (GME) estimator: formulation and a monte carlo study // VII. National Symposium on Econometrics and Statistics, Istanbul, Turkey, May 26-27, 2005.
4. *Ximing Wu.* A Weighted Generalized Maximum Entropy Estimator with a Data-driven Weight // Entropy, 2009. 11. 1-x.
5. *Попков Ю. С., Дубнов Ю. А.* Энтропийно-робастное рандомизированное прогнозирование при малых объемах ретроспективных данных // Автоматика и телемеханика, 2015. (в печати)
6. *Магнус Я. Р., Катыйшев П. К., Пересецкий А. А.* Эконометрика. Начальный курс: Учеб. – 6-е изд., перераб. и доп. – М.: Дело, 2004. – 576 с.
7. *А. Зельнер.* Байесовские методы в эконометрии. // Пер. с англ. Г. Г. Пирогова и Ю. П. Федоровского – М.: Статистика, 1980. – 438 с.

# Исследование среднеквадратичного риска при обработке звуковых сигналов FDR-порогом

А. Ю. Заспа<sup>72</sup>

**Аннотация:** В данной работе рассматривается эффективность применения теории множественной проверки гипотез в задаче обработки сигналов на примере FDR-порога. Показывается возможность использования FDR-порога при пороговой фильтрации, используя метод Бенджамини-Хочберга множественной проверки гипотез. Эффективность метода проверяется за счет сравнения с другими наиболее популярными порогами, такими как универсальный и SURE-порог.

**Ключевые слова:** FDR-порог; вейвлет-анализ; пороговая обработка

## Введение

В жизни мы часто встречаемся с задачей обработки цифровых сигналов, это могут быть сигналы совершенно разных видов и происхождения. Наиболее распространенной задачей из этой области является фильтрация сигналов, так как она чаще всего становится отправной точкой для последующей обработки сигнала.

Построение алгоритмов цифровой фильтрации опирается на некоторые предпосылки о вероятностной модели сигнала и шума, а также на статистические критерии оптимальности. Всевозможные предположения о вероятностном распределении шума и различные статистические критерии порождают множество различных методов цифровой обработки.

---

<sup>72</sup>ВМК МГУ им. М. В. Ломоносова; 1994 г.р.; [zaspa@ya.ru](mailto:zaspa@ya.ru)

# 1. Основы вейвлет-анализа

## 1.1. Кратномасштабное вейвлет-представление сигнала

Для начала необходимо получить коэффициенты вейвлет-разложения сигнала. Это возможно сделать за счет кратномасштабного анализа (набора вложенных подпространств  $V_j$  в  $L^2(\mathbb{R})$  и ортогональных дополнений  $W_j$  подпространств  $V_j$  до подпространств  $V_{j+1}$ ) [1]. Зафиксируем некоторое значение уровня аппроксимации  $j_0$  и рассмотрим ортогональную проекцию сигнала на подпространство  $V_{j_0}$ . Оператор проектирования  $P_{j_0}(f)$  будет иметь вид:

$$P_{j_0}(f) = \sum_{k \in \mathbb{Z}} \langle f(x), \phi_{j_0,k}(x) \rangle \phi_{j_0,k}(x) = \sum_{k \in \mathbb{Z}} a_{j_0,k} \phi_{j_0,k}(x).$$

Соответственно, представление сигнала в базисах подпространствах  $V_{j_0-1}$  и  $W_{j_0-1}$  выглядит так:

$$P_{j_0}(f) = \sum_{k \in \mathbb{Z}} a_{j_0-1,k} \phi_{j_0-1,k}(x) + \sum_{k \in \mathbb{Z}} d_{j_0-1,k} \psi_{j_0-1,k}(x),$$

где  $a_{j_0-1,k}$  — соответственно аппроксимирующие коэффициенты  $(j_0 - 1)$ -го уровня, а  $d_{j_0-1,k}$  называются *детализирующими коэффициентами*  $(j_0 - 1)$ -го уровня.

$$a_{j,k} = \int_{-\infty}^{+\infty} f(x) \phi_{j,k}(x) dx \quad (1)$$

$$d_{j,k} = \int_{-\infty}^{+\infty} f(x) \psi_{j,k}(x) dx \quad (2)$$

*Вейвлет-представлением* (или *вейвлет-разложением*) функции  $f(x)$  с глубиной разложения  $J$  будем называть запись:

$$\widehat{f}(x) = \sum_{k \in \mathbb{Z}} a_{j_0-J,k} \phi_{j_0-J,k}(x) + \sum_{j=1}^J \sum_{k \in \mathbb{Z}} d_{j_0-j,k} \psi_{j_0-j,k}(x).$$

## 2. Пороговая обработка

### 2.1. Модель аддитивного шума

Одним из главных направлений цифровой обработки сигналов является очистка сигнала от шумовых компонент. Чаще всего можно принять гипотезу о случайности шума, и что случайные величины, соответствующие шуму в каждом отсчете функции, независимы и одинаково распределены. В данной модели принимается предположение о том, что шум имеет нормальное распределение. Также логично будет учесть, что помехи сосредоточены в высокочастотной области спектра, а исходный («незашумленный») сигнал - в низкочастотной.

Запишем сигнал следующим образом:

$$\tilde{f}_i = f(x_i) + \varepsilon_i = f_i + \varepsilon_i, \quad i = 1, \dots, N, \quad (3)$$

где  $f_i = f(x_i)$  – точные значения сигнала в узлах равномерной сетки;  $\varepsilon_i$  – случайная величина, соответствующая шумовой компоненте измерения  $f(x_i)$ .

### 2.2. Пороговая фильтрация

Для таких предположений можно составить следующий алгоритм удаления шумовой составляющей: разложение сигнала по базису вейвлетов, выбор порогового значения шума на каждом уровне разложения, сама пороговая фильтрация коэффициентов, реконструкция (обратное вейвлет-преобразование) сигнала. Коэффициенты вейвлет-разложения (1) и (2), которые меньше заданного порога, зануляются, те, что больше – либо остаются прежними, либо из них вычитается порог.

## 3. FDR-порог

### 3.1. Множественная проверка гипотез

Одним из методов построения порогового значения является использование алгоритмов множественной проверки гипотез. Так как  $\tilde{f}_i$  являются реализациями некоторых случайных величин (3), то можно каждому зашумленному коэффициенту  $\tilde{f}_i$ ,  $i =$

$1, \dots, N$  сопоставить нулевые гипотезы  $\{H_{0_i}, i = 1, \dots, N\}$ . Каждая нулевая гипотеза включает в себя предположение о распределении этой случайной величины и о наличии шумовой компоненты в коэффициенте.

Существует несколько мер, оценивающих число произошедших ошибок первого рода, при использовании определенного алгоритма проверки гипотез. Одной из таких мер является **FDR**. FDR (англ. *False Discovery Rate*) – средняя доля ошибок первого рода среди всех отклоненных гипотез:

$$\mathbf{FDR} = \mathbb{E} \frac{V}{R}, \text{ при } R > 0.$$

### 3.2. Процедура Бенджамини-Хочберга контроля FDR

В своей работе [2] Бенджамини и Хочберг доказали, что приведенный ими алгоритм множественной проверки гипотез позволяет ограничить сверху значение FDR на уровне  $q^*$ , т.е.

$$\mathbb{E} \frac{V}{k} \leq q^*,$$

где  $V$  – количество ошибок первого рода.

Опишем саму процедуру:

В качестве нулевых гипотез возьмем гипотезы, что коэффициенты являются полностью шумовыми и, в соответствии с используемой моделью шума, имеют нормальное распределение с нулевым математическим ожиданием.

Пусть  $H_0 = \{H_{0_1}, H_{0_2}, \dots, H_{0_N}\}$  проверяемые нулевые гипотезы с соответствующими  $p$ -значениями  $p = \{p_1, p_2, \dots, p_N\}$ . Тогда упорядочив по возрастанию эти  $p$ -значения, получим набор  $p' = \{p_{(1)}, p_{(2)}, \dots, p_{(N)}\}$ , и сопоставим каждому  $p$ -значению  $p_{(i)}$  соответствующую гипотезу  $H_{0_{(i)}}$ . В итоге получим набор гипотез  $H'_0 = \{H_{0_{(1)}}, H_{0_{(2)}}, \dots, H_{0_{(N)}}\}$ .

возьмем за  $k$  такое максимальное  $i$  для которого выполняется

$$p_{(i)} \leq \frac{i}{N} q^*;$$

затем отвергнем все нулевые гипотезы  $H_{0_{(i)}}$  с номерами  $i = 1, 2, \dots, k$ .

Так как в качестве  $H_0$  взято предположение о нормальном распределении коэффициентов, то р-значения будут выражаться следующим образом:

$$p_i = 2 \left[ 1 - \Phi \left( \frac{|\tilde{f}_i|}{\sigma} \right) \right], \quad (4)$$

где  $\Phi(x)$  – функция стандартного нормального распределения  $N(0, 1)$ .

### 3.3. Полученные результаты

Вначале проверим эффективность работы данного порога «на слух». Для этого возьмем звуковую дорожку, имеющую ярко выраженную шумовую компоненту и обработаем FDR-порогом. Помехи в сигнале стали практически не слышны, т.е. фоновая составляющая была удалена с хорошей точностью. Но весь звук дорожки стал «глухой», из-за того, что стали преобладать низкочастотные компоненты сигнала. В результате можно сказать, что FDR-порог хорошо работает на сигналах, имеющих постоянную фоновую шумовую составляющую, в случаях когда надо полностью избавиться от помех, допуская при этом потерю некоторых информативных коэффициентов.

Для исследования среднеквадратичного риска будем сравнивать экспериментальные значения риска, полученные по известному «чистому» сигналу. Сравним с рисками при обработке универсальным и SURE-порогом. Универсальный порог ( $T = \sigma\sqrt{2\ln n}$ , где  $n$  - кол-во отсчетов, а  $\sigma$  - дисперсия шума) был предложен Донохо и Джонстоном, основываясь на свойствах нормального распределения. В свою очередь, SURE-порог вычисляется как порог, минимизирующий несмещенную оценку риска SURE. Эта оценка имеет вид:  $\hat{r}_t(f, T) = \sum_{i=1}^N \Phi(\tilde{f}_i^2)$ , где  $\Phi(x) =$

$\begin{cases} x - \sigma^2, & \text{если } x \leq T^2; \\ \sigma^2 + T^2, & \text{если } x > T^2. \end{cases}$  Для этого будем зашумлять сигнал «белым» шумом с заданной дисперсией, генерируя данное распределение с помощью псевдослучайных чисел. Для После этого разложим полученный сигнал быстрым вейвлет-преобразованием с вейвлетом Добеши. [3]. В качестве тестовых сигналов возьмем два звуковых

сигнала. Первый содержит звук музыкальных инструментов, второй – голос человека. Будем вычислять риски при разных уровнях отношения шум/сигнал. Для демонстрации результатов приводится выдержка из таблицы для первого сигнала «Таблица 1».

**Таблица 1.** Значения риска обработки для первого сигнала

Отношение шум/сигнал	Пороги		
	FDR	Универсальный	SURE
0.0747	581.9	1149.1	453.1
	574.8	1174.6	449.2
	582.9	1156.2	451.6
	585.8	1149.1	452.5
	584.8	1157.8	450.5
0.3689	1728.6	3366.7	1318.5
	1724.1	3281.5	1325.7
	1725.9	3293.4	1329
	1742.3	3282.8	1324.6
	1745.8	3338.2	1332.5
0.83	3102.8	5055.4	2205.9
	3116.4	4997.3	2240.5
	3111.8	5018.8	2223.5
	3121.4	5037.9	2234.8
	3109.5	4963.7	2236.1

По полученным результатам можно сказать, что фильтрация с FDR-порогом дает средние значения риска. Риск FDR-обработки примерно в полтора раза ниже риска универсального порога и на 30% хуже, чем риск при фильтрации SURE-порогом. При увеличении отношения шум/сигнал квадратичный риск FDR-порога немного отдаляется от SURE и приближается к универсальному. Но после того, как отношение шум/сигнал становится больше 1, квадратичный риск FDR-порога снова начинает сближаться с риском SURE-порога. Это возникает из-за того, что SURE-порог плохо приспособлен для обработки сигналов, где энергия чистого сигнала меньше энергии шума.

## Выводы

В ходе работы была показана возможность построения порога фильтрации вейвлет-коэффициентов методом Бенджамини-Хочберга. На нескольких реальных звуковых сигналах была протестирована обработка с данным порогом. В полученном сигнале шумовую компоненту слышно не было, но звук стал «глухим», т.е. были потеряны высокочастотные составляющие. С помощью программы были получены значения риска для разных порогов в зависимости от отношения шум/сигнал. По полученным экспериментальным данным были сделаны выводы о соотношении среднеквадратичного риска между порогами, выявлены зависимости риска от отношения шум/сигнал. Среднеквадратичный риск FDR-порога оказался значительно лучше риска универсального порога и немного хуже риска при обработке SURE-порогом. По полученным результатам можно сделать вывод, что FDR-порог хорошо подходит для фильтрации сильно зашумленных сигналов, когда гораздо важнее убрать весь шум, чем потерять несколько информативных коэффициентов.

## Благодарности

*Хочу поблагодарить своего научного руководителя доцента О.В. Шестакова за помощь в поиске материала, за разъяснение принципов работы различных алгоритмов.*

## Список литературы

1. *Захарова Т. В., Шестаков О. В.* Вейвлет-анализ и его приложения. – М.: МАКС Пресс, 2009. – 152 с.
2. *Benjamini Y., Hochberg Y.* Controlling the false discovery rate: a practical and powerful approach to multiple testing // *Journal of the Roy Statist. Soc. Ser. B.*, 1995. No.57. С. 289–300.
3. *Добеши И.* Десять лекций по вейвлетам. – И.: НИЦ "Регулярная и хаотическая динамика 2001. – 464 с.

# Имитационная модель для оценки интерференции в одноранговых сетях беспроводных устройств с прямоугольными кластерами \*

Р. Н. Ковальчуков<sup>73</sup>, Ю. В. Гайдамака<sup>74</sup>

**Аннотация:** При реализации новой концепции взаимодействия беспроводных устройств (D2D, device-to-device соединения) в современных сетях активно используется принцип повторного использования радиоресурсов. В этом случае при оценке интерференции следует учитывать особенности D2D-соединений, которые требуют разработки специальной методики. В статье описан интерфейс разработанного программного средства для имитационного моделирования случайной величины отношения сигнала к интерференции (*SIR*, Signal-to-Interference Ratio) в беспроводных сетях, зона обслуживания которых разбита на кластеры прямоугольной формы. В качестве иллюстрации работы имитационной модели для наборов исходных данных построены графики функции плотности распределения СВ *SIR*, а также приведены значения числовых характеристик этой СВ.

**Ключевые слова:** SIR; беспроводные сети; имитационное моделирование; D2D

## Введение

В статье модифицирована имитационная модель для анализа показателей, влияющих на качество предоставления услуг в одноранговых сетях беспроводной связи с повторным использованием радиочастотных ресурсов, зоны обслуживания которых имеют прямоугольную форму, разработанная в [1]. Рассмотрим модель фрагмента беспроводной сети, где устройства взаимодействуют в 9 смежных зонах прямоугольной формы (Рис. 1), далее назы-

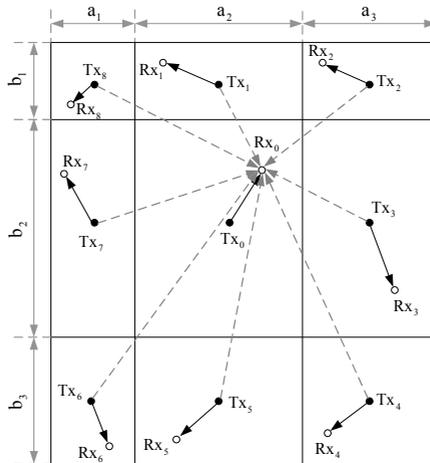
---

\*Работа выполнена при финансовой поддержке РФФИ, проекты № 14-07-00090, 15-07-03051.

<sup>73</sup>РУДН; 1993 г.р.; rnkovalchukov@sci.pfu.edu.ru

<sup>74</sup>РУДН; к.ф.-м.н, доцент; ygaidamaka@sci.pfu.edu.ru

ваемых кластерами. Конфигурацию, подобную рассматриваемой, можно встретить внутри офисных, жилых или многофункциональных зданий в условиях городской местности. Модель для случая двух соседних зон прямоугольной формы была рассмотрена в [2].



**Рис. 1.** Схема взаимодействия устройств в прямоугольных кластерах

В каждой зоне произвольно расположена пара устройств, каждое из которых может выступать как в качестве приемника, так и в качестве передатчика. Согласно принципу повторного распределения радиочастот для обслуживания соединения между устройствами в соседних зонах могут быть использованы одни и те же частотные ресурсы, например, ресурсные блоки (PRB, physical resource block) в сетях LTE (Long Term Evolution). В этом случае пары из соседних кластеров будут создавать помехи друг другу.

Задача состоит в оценке отношения сигнала к интерференции плюс шум (SINR, Signal to Interference plus Noise Ratio), которое является одной из основных характеристик качества канала в беспроводных сетях связи:

$$SINR = \frac{S}{\sum_{i=1}^N I_i + \sigma^2},$$

где  $S$  – мощность сигнала,  $N$  – число интерферирующих объектов,  $I_i$  – мощность интерферирующего сигнала от  $i$ -го объекта,  $\sigma^2$  – мощность шума.

Мощности сигналов  $S$  и  $I_i$  являются функциями от расстояния между рассматриваемым приемником и передатчиком или интерферирующим объектом и могут быть вычислены по следующим формулам:

$$S := S(R_0) = g_0 R_0^{-\alpha_0},$$

$$I_i := I_i(D_i) = g_i D_i^{-\alpha_i}, i = 1, \dots, N,$$

где  $g_i$  – базовая мощность сигнала на передатчике,  $\alpha_i$  – коэффициент потерь, т.н. «экспонента затухания», принимающий значение от 2 (при условии прямой видимости) до 6 (в худшем случае),  $i = 0, \dots, N$ .

Для рассматриваемого случая  $N = 8$  соседних кластеров при упрощающем предположении об отсутствии шума ( $\sigma^2 = 0$ ) и равенстве мощностей сигналов на передатчиках  $g_i = g$  будем оценивать отношение сигнала к интерференции ( $SIR$ , Signal to Interference Ratio):

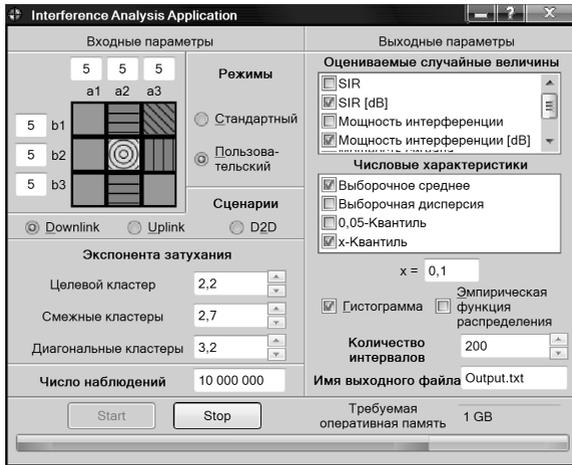
$$SIR = \frac{S(R_0)}{\sum_{i=1}^8 I(D_i)}.$$

Разработанная имитационная модель на основе метода Монте-Карло представляет собою программное средство на языке C++, которое реализует сбор статистики при моделировании случайной величины  $SIR$  для приемника центрального кластера.

## 1. Интерфейс симулятора

Графический пользовательский интерфейс для семейства ОС Windows показан на Рис. 2. В левой части окна определяются входные параметры моделирования. Ими являются длины сторон прямоугольных зон обслуживания и значения экспонент затухания для рассматриваемых кластеров. Оценка величины  $SIR$  производится либо в т.н. «стандартном» режиме, когда учитывается влияние интерферирующих сигналов из всех 8 соседних кластеров, либо в т.н. «пользовательском» режиме, когда пользователь симулятора

вручную выбирает набор кластеров, интерферирующий сигнал из которых будет учитываться в модели. На Рис. 2 показан пример выбора пользовательского режима. Здесь центральный кластер, для которого производится оценка величины  $SIR$ , отмечен концентрическими кругами, интерферирующие кластеры обозначены прямыми линиями, а кластеры, которые не участвуют в симуляции, оставлены пустыми.



**Рис. 2.** Графический интерфейс пользователя симулятора «Interference Analysis Application»

Имитационная модель предусматривает три варианта расположения устройств внутри зоны [3]: т.н. «Uplink» сценарий, когда передатчики находятся в центрах своих прямоугольных зон, а координаты приемников имеют равномерное распределение, т.н. «Downlink» сценарий, когда приемники находятся в центрах своих зон, а координаты передатчиков имеют равномерное распределение, т.н. «D2D» сценарий, когда координаты и приемников, и передатчиков имеют равномерное распределение.

В следующем разделе выбираются значения экспонент затухания  $\alpha_i$ , где  $i = 0$  для целевого кластера,  $i = 1, 3, 5, 7$  для соседних сверху, снизу, слева и справа кластеров,  $i = 2, 4, 6, 9$  для смежных по диагонали кластеров. Последним входным параметром является «Число наблюдений» – количество розыгрышей случай-

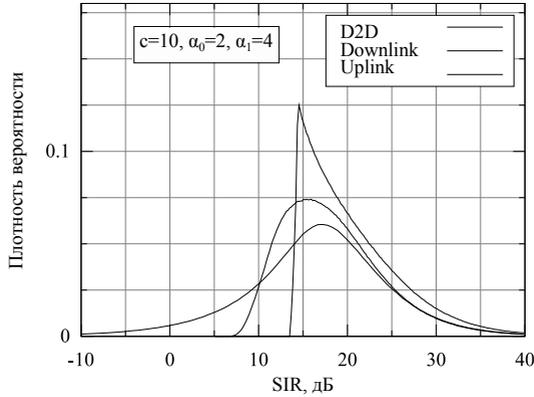
ных величин для сбора статистики.

В правой части окна задаются выходные параметры симулятора. Пользователь может выбрать один или несколько показателей для оценки, в т.ч. *SIR*, мощность полезного сигнала, мощность интерферирующего сигнала, которые являются случайными величинами и будут смоделированы при запуске программы, а также числовые характеристики интересующего показателя. При анализе интерференции отдельный интерес представляет квантиль уровня 0,05, которую используют для оценки качества беспроводной связи, поэтому эта характеристика вынесена в отдельную строку. Вид выводимых результатов можно выбрать, поставив флажок у параметра «Гистограмма» или «Эмпирическая функция распределения». Два последних параметра отвечают за количество интервалов гистограммы и эмпирической функции распределения случайных величин и имя выходного файла, в который будут записаны результаты работы программы.

Для запуска симуляции после выбора всех параметров нужно нажать на кнопку «Start». При необходимости выполнение можно прервать кнопкой «Stop». Также в окне выводится информация о минимально необходимом объеме свободной оперативной памяти для работы программы и шкала, отражающая ход выполнения программы.

## 2. Численный эксперимент

Для численного анализа взаимодействия одноранговых устройств в смежных зонах покрытия прямоугольной формы, в соответствии со схемой, показанной на Рис. 1, с помощью разработанного программного обеспечения, был проведен численный эксперимент по оценке характеристик случайной величины *SIR* – математического ожидания, среднеквадратического отклонения и квантилей. В имитационной модели учитывается, что экспоненты затухания  $\alpha$  в центральной и смежных ячейках зависят от расположения ячеек следующим образом:  $\alpha_0 = 2$  для центральной ячейки,  $\alpha_1 = \alpha_3 = \alpha_5 = \alpha_7 = 4$  для ячеек, смежных с центральной. Устройства, расположенные по диагонали относительно центрального кластера, не принимали участия в эксперименте, в обратном случае нужно выбирать  $\alpha_2 = \alpha_4 = \alpha_6 = \alpha_8 > \alpha_1$ .



**Рис. 3.** Оценка плотности распределения случайной величины  $SIR$

На Рис. 3 показана эмпирическая плотность случайной величины  $SIR$  в диапазоне значений от  $-10$  до  $40$  дБ для серии из  $10^7$  экспериментов. Оценка характеристик приведена в Таблице 1. При моделировании предполагалось, что все ячейки, в которых находятся взаимодействующие устройства, имеют одинаковые размеры сторон  $a_i = b_i = c = 10$ . Заметим, что результаты, полученные в [3], были повторены с использованием симулятора «Interference Analysis Application».

**Таблица 1.** Оценка характеристик с.в.  $SIR$ , [дБ]

	<b>Downlink</b>	<b>Uplink</b>	<b>D2D</b>
$M[SIR]$	20, 64	21, 33	16, 55
$\sigma[SIR]$	5, 66	5, 33	8, 52
$SIR_{0,05}$	14, 46	14, 19	1, 95

## Выводы

Оценка интерференции в современных сетях взаимодействия беспроводных устройств представляет собою сложную задачу. Для случая кластеров прямоугольной формы аналитические

формулы для плотности распределения величины  $SIR$  могут быть получены лишь для частных случаев, например, для «Uplink» и «Downlink» сценариев и простых видов распределения устройств в кластерах. Оценка величины  $SIR$  в случае, когда и приемник, и передатчик, занимают случайные позиции в прямоугольнике (т.н. «D2D» сценарий), возможна лишь при помощи имитационного моделирования. Заметим, что для случая кластеров круглой формы получение замкнутых аналитических формул затруднительно, поэтому для оценки величины  $SIR$  следует использовать приближенные методы или имитационное моделирование [4]. Задачей дальнейших исследований является расширение симулятора для трехмерного случая, а также исследование дополнительных моделей затухания сигнала.

## Список литературы

1. *Гайдамака Ю. В., Самуйлов А. К., Бегишев В. О., Ковальчуков Р. Н.* Имитационная модель для оценки интерференции при взаимодействии беспроводных устройств в прямоугольных кластерах // Тр. 9-й междунар. научн.-техн. отрасл. конф. «Технологии информационного общества» – М.: МГУСИ, 2015. С. 7.
2. *Самуйлов А. К., Бегишев В. О., Ковальчуков Р. Н.* Расчет интерференции при взаимодействии беспроводных устройств в квадратных кластерах // Тр. всеросс. конф. ИТТММ. – М.: РУДН, 2015. С. 110.
3. *Begishev V., Kovalchukov R., Samuylov A., Ometov A., Moltchanov D., Gaidamaka Yu., Andreev S., Koucheryavy E.* An analytical approach for SINR estimation in adjacent rectangular premises // LNCS 9247, 2015. Pp. 446–458.
4. *Самуйлов А. К., Гайдамака Ю. В.* Метод расчета характеристик интерференции двух взаимодействующих устройств в беспроводной гетерогенной сети // Информатика и ее применения, 2015. Т.9. No.1. С. 9–14.

# Эргодичность и устойчивость системы обслуживания с одним сервером\*

А. В. Коротышева<sup>75</sup>, К. М. Киселева<sup>76</sup>, Я. А. Сатин<sup>77</sup>

**Аннотация:** Рассматривается система массового обслуживания одного сервера. Новые клиенты приходят в соответствии с процессом Пуассона со скоростью  $\lambda(t)$ . Клиенты имеют независимое и одинаково распределенное экспоненциальное время обслуживания  $\{S_i\}$  с общим элементом  $S$  и скоростью  $\mu(t) := 1/ES$ . Если новый клиент находит сервер занятым, он присоединяется на орбиту бесконечного потенциала и называется вторичным клиентом. Мы предполагаем, что орбита работает как одиночный сервер обслуживания в порядке поступления, то есть если орбита непустая, то прежний вторичный клиент пытается попасть на сервер согласно экспоненциальному распределению со скоростью  $\mu_0(t)$ . Для получения оценок эргодичности и устойчивости мы применяем общий подход, в котором используем логарифмическую норму матрицы.

**Ключевые слова:** система обслуживания; экспоненциальная эргодичность; нуль-эргодичность; устойчивость.

## Введение

Для текущего  $t$  обозначим через  $\nu(t)$  число клиентов сервера, а через  $N(t)$  общее количество клиентов на орбите. То есть,  $\nu(t) = 0$ , если сервер пуст,  $\nu(t) = 1$  – в противном случае, при  $N(t) = 0, 1, \dots$ . Введем основной двумерный процесс  $X(t) = \{\nu(t), N(t), t \geq 0\}$  с пространством состояний  $\{0, 1\} \times \{0, 1, 2, \dots\}$  (см. [1]). Рассмотрим переходы между состояниями системы. Для начала, перечислим состояния процесса следующим образом: каждое состояние  $\{0, n\}$  обозначим через  $2n + 1$  при  $n \geq 0$ , а каждое состояние  $\{1, n\}$  –

---

\* Исследование выполнено за счет гранта Российского научного фонда (проект № 14-11-00397)

<sup>75</sup>ВоГУ, ФИЦ ИУ РАН (ИПИ РАН); 1988 г.р.; к. ф.-м. н.; a\_korotyшева@mail.ru

<sup>76</sup>ВоГУ, ФИЦ ИУ РАН (ИПИ РАН); 1992 г.р.; ksushakiseleva@mail.ru

<sup>77</sup>ВоГУ, ФИЦ ИУ РАН (ИПИ РАН); 1978 г.р.; к. ф.-м. н.; yacovi@mail.ru

через  $2n, n \geq 1$ . Соответствующую данному перечислению матрицу интенсивностей обозначим через  $Q = (q_{ij})$ . Отсюда получаем, что

$$q_{1,1} = -\lambda(t), \quad q_{1,2} = \lambda(t),$$

и для  $n \geq 1$ ,

$$q_{2n,2n} = -(\lambda(t) + \mu(t)), \quad q_{2n,2n+2} = \lambda(t), \quad q_{2n,2n-1} = \mu(t), \\ q_{2n+1,2n+1} = -(\lambda(t) + \mu_0(t)), \quad q_{2n+1,2n} = \mu_0(t), \quad q_{2n+1,2n+2} = \lambda(t).$$

Таким образом, матрица интенсивностей  $Q = (q_{ij})$  выглядит так

$$Q(t) = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ \mu & -(\lambda + \mu) & 0 & \lambda & 0 & 0 & 0 & 0 & \dots \\ 0 & \mu_0 & -(\lambda + \mu_0) & \lambda & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & \mu & -(\lambda + \mu) & 0 & \lambda & 0 & 0 & \dots \\ 0 & 0 & 0 & \mu_0 & -(\lambda + \mu_0) & \lambda & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \mu & -(\lambda + \mu) & 0 & \lambda & \dots \\ \vdots & \ddots \end{pmatrix}.$$

Пусть  $p_{ij}(s, t) = Pr \{X(t) = j | X(s) = i\}$ ,  $i, j \geq 1$ ,  $0 \leq s \leq t$ , переходные вероятности процесса  $X = X(t)$ , а  $p_i(t) = Pr \{X(t) = i\}$  – его вероятности состояний.

Вероятностная динамика процесса описывается системой Колмогорова:

$$\frac{d\mathbf{p}}{dt} = A\mathbf{p}$$

где  $A = (a_{ij})_{i,j=1}^{\infty} = (q_{ji})_{i,j=1}^{\infty} = Q^T$  – соответствующая транспонированная матрица интенсивностей, а  $\mathbf{p} = \mathbf{p}(t) = (p_1(t), p_2(t), \dots)^T$  – вектор-столбец вероятностей состояний процесса  $X(t)$ .

Будем предполагать, что функции  $\lambda(t)$ ,  $\mu(t)$  и  $\mu_0(t)$  локально интегрируемы на  $[0; \infty)$ , и ограничены, причем

$$\lambda(t) \leq \Lambda, \quad \mu(t) \leq \Gamma, \quad \mu_0(t) \leq \Gamma, \quad t \geq 0.$$

Далее через  $\|\cdot\|$  будем обозначать  $l_1$ -норму, т.е.  $\|\mathbf{x}\| = \sum |x_i|$ , а  $\|B\| = \sup_j \sum_i |b_{ij}|$  для матрицы  $B$ .

# 1. Нуль-эргодичность

Рассмотрим убывающую последовательность положительных чисел  $\{\delta_i\}$ ,  $i = 1, 2, \dots$ ,  $\delta_1 = 1$ , и соответствующую диагональную матрицу  $\Delta$ :

$$\Delta = \begin{pmatrix} \delta_1 & 0 & 0 & \dots \\ 0 & \delta_2 & 0 & \dots \\ 0 & 0 & \delta_3 & \dots \\ \dots & & & \dots \end{pmatrix}$$

Пусть  $l_{1\Delta} = \{\mathbf{p} = (p_1, p_2, \dots)^T : \|\mathbf{p}\|_{1\Delta} \equiv \|\Delta \mathbf{p}\| < \infty\}$  - пространство последовательностей:

Положим  $\delta_{2k+1} = a\delta_{2k}$ ,  $\delta_{2k} = b\delta_{2k-1}$  при  $k \geq 1$  для некоторых  $a < 1$ ,  $b < 1$ .

Найдем логарифмическую норму матрицы интенсивностей в новом пространстве:

$$\begin{aligned} \gamma(A(t))_{1\Delta} &= \gamma(\Delta A(t) \Delta^{-1})_1 = \sup_i \left( a_{ii}(t) + \sum_{j \neq i} \frac{\delta_j}{\delta_i} a_{ij}(t) \right) = \\ &= -\min(\lambda(t)(1-b), \lambda(t)(1-ab) - \mu(t)(b^{-1}-1), \\ &\quad \lambda(t)(1-b) - \mu_0(t)(a^{-1}-1)). \end{aligned}$$

Пусть  $b$  из интервала  $\left(\frac{(c+1)d}{cd+c+1}, 1\right)$ , где  $c = \inf_t \frac{\lambda(t)}{\mu_0(t)}$ ,  $d = \sup_t \frac{\mu(t)}{\lambda(t)}$ . Тогда существует постоянное  $a$ , при условии, что  $\sup_t \frac{\mu(t)}{\lambda(t)} < \inf_t \frac{\lambda(t)}{\mu_0(t)} + 1$ :

$$\frac{1}{\inf_t \frac{\lambda(t)}{\mu_0(t)} (1-b) + 1} < a < \frac{1}{b} - \sup_t \frac{\mu(t)}{\lambda(t)} \left( \frac{1}{b^2} - \frac{1}{b} \right).$$

Таким образом, справедлива следующая теорема.

**Теорема 1.** Пусть  $\sup_t \frac{\mu(t)}{\lambda(t)} < \inf_t \frac{\lambda(t)}{\mu_0(t)} + 1$ . Кроме того,  $\int_0^t \zeta(\tau) d\tau \rightarrow \infty$  при  $t \rightarrow \infty$ . Тогда процесс  $X(t)$  нуль-эргодичен и справедливо следующее неравенство

$$\sum_{i=0}^N p_i(t) \leq \frac{\delta_k}{\delta_N} e^{-\int_0^t \zeta(\tau) d\tau}$$

для любого  $t \geq 0$ , начального условия  $X(0) = k$ , и любого натурального  $N$ , где

$$\zeta(\tau) = \min(\lambda(\tau)(1-ab) - \mu(\tau)(b^{-1} - 1), \\ \lambda(\tau)(1-b) - \mu_0(\tau)(a^{-1} - 1)) > 0.$$

## 2. Экспоненциальная эргодичность

Пусть  $p_1(t) = 1 - \sum_{i \geq 2} p_i(t)$  (подробнее см. [3]). Тогда имеем следующее равенство

$$\frac{d\mathbf{z}(t)}{dt} = B(t)\mathbf{z}(t) + \mathbf{f}(t),$$

где  $\mathbf{f}(t) = (a_{20}(t), a_{30}(t), \dots)^T = (\lambda(t), 0, 0, \dots)^T$ ,  $\mathbf{z}(t) = (p_2(t), p_3(t), \dots)^T$ ,

$$B(t) = (b_{ij}(t))_{i,j=1}^{\infty} = \begin{pmatrix} a_{22} - a_{20} & a_{23} - a_{20} & \cdots & a_{2r} - a_{30} & \cdots \\ a_{32} - a_{30} & a_{33} - a_{30} & \cdots & a_{3r} - a_{30} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{r1} - a_{r0} & a_{r2} - a_{r0} & \cdots & a_{rr} - a_{r0} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

Рассмотрим последовательность положительных чисел  $\{d_i\}$ ,  $i = 2, 3, \dots$ ,  $d_2 = 1$ . Положим  $g_i = \sum_{n=2}^i d_n$ . Пусть  $D$  - верхняя треугольная матрица,

$$D = \begin{pmatrix} d_2 & d_2 & d_2 & \cdots \\ 0 & d_3 & d_3 & \cdots \\ 0 & 0 & d_4 & \cdots \\ & \ddots & \ddots & \ddots \end{pmatrix}, \quad (1)$$

а  $l_{1D}$  соответствующее пространство последовательностей

$$l_{1D} = \{\mathbf{z} = (p_2, p_3, \dots)^T \mid \|\mathbf{z}\|_{1D} \equiv \|\mathbf{Dz}\|_1 < \infty\}.$$

Положим  $d_{2k+1} = ad_{2k}$ ,  $d_{2k} = bd_{2k-1}$  при  $k \geq 1$ . Тогда

$$\gamma(B(t))_{1D} = \gamma(DB(t)D^{-1})_1 = -\min(\lambda(t) + \mu(t) - \lambda(t)(b+ab), \\ \lambda(t) + \mu_0(t) - \mu(t)b^{-1}, \lambda(t) + \mu(t) - \lambda(t)(b+ab) - \mu_0(t)a^{-1}).$$

Положим  $x = ab$  и рассмотрим неравенство:

$$\frac{\mu(t)}{\lambda(t) + \mu_0(t)} < b < \frac{\lambda(t) + \mu(t) - x\lambda(t)}{\lambda(t) + \mu_0(t)/x}. \quad (2)$$

Для удобства обозначим  $c = \inf_t \frac{\mu(t)}{\lambda(t)}$ ,  $d = \inf_t \frac{\lambda(t)}{\mu(t)}$ . Неравенство (2) справедливо для некоторого  $b$  и  $a = x/b$ , если  $x \in (1; \frac{c}{d(c+1)})$ .

**Теорема 2.** Пусть  $\sup_t \frac{\mu(t)}{\lambda(t)} > \sup_t \frac{\lambda(t)}{\mu_0(t)} + 1$ . Кроме того,  $\int_0^t \alpha^*(\tau) d\tau \rightarrow \infty$  при  $t \rightarrow \infty$ . Тогда процесс  $X(t)$  экспоненциально эргодичен и выполняется следующее неравенство:

$$\| \mathbf{p}^*(t) - \mathbf{p}^{**}(t) \| \leq 4e^{-\int_0^t \alpha^*(\tau) d\tau} \sum_{i \geq 1} g_i | \mathbf{p}^*(0) - \mathbf{p}^{**}(0) |$$

для любого  $t \geq 0$  и начальных состояний  $\mathbf{p}^*(0)$ ,  $\mathbf{p}^{**}(0)$ , где

$$\alpha^* = \min (\lambda(t) + \mu_0(t) - \mu(t)b^{-1}, \lambda(t) + \mu(t) - \lambda(t)(b + ab) - \mu_0(t)a^{-1}).$$

### 3. Оценки устойчивости

Рассмотрим теперь «возмущенный» процесс  $\bar{X} = \bar{X}(t)$ ,  $t \geq 0$ , обозначая все его характеристики с чертой (см. [2]). Введем обозначение  $\hat{A}(t) = A - \bar{A}$ . Для простоты записи оценок будем предполагать, что возмущения «равномерно малы», то есть при всех  $n$  и почти всех  $t \geq 0$  выполняется неравенство  $\|\hat{A}(t)\| \leq \varepsilon$ .

Рассмотрим также последовательность положительных чисел  $\{d_i\}$ ,  $i = 1, 2, \dots$ ,  $d_1 = 1$  и матрицу  $D$  (см. (1)).

Положим

$$\alpha_i(t) = \lambda(t) \left( 1 - \frac{d_{i+1} + d_{i+2}}{d_i} \right) - \frac{d_i}{d_{i+1}} \mu(t) - \frac{d_{i-1}}{d_i} \mu_0(t), \quad i \geq 1,$$

где  $d_0 = 0$ . Пусть  $d = \min_{i \geq 1} d_i$ ,  $W = \min_{i \geq 1} \frac{d_i}{i}$ .

**Теорема 3.** Пусть существует последовательность положительных чисел  $\{d_j\}$  такая, что существует положительное число  $\theta^* \leq \int_0^1 \theta(t) dt$ , где  $\theta(t) \leq \alpha_i(t)$  и  $\theta(t) > 0$ ,  $i = 1, 2, \dots$ ,  $t \geq 0$ . Пусть также  $\lambda(t)$ ,  $\mu(t)$  и  $\mu_0(t)$  1-периодичны и

$$K = \sup_{|t-s| \leq 1} \int_s^t \theta(\tau) d\tau < \infty.$$

Тогда для любых начальных условий  $p(0)$  и  $\bar{p}(0)$  процессов  $X(t)$  и  $\bar{X}(t)$  соответственно справедливы неравенства

$$|p(t) - \bar{p}(t)| \leq \frac{4e^K \varepsilon e^{\left(1 + \max_{|i-j|=1} \frac{d_i}{d_j}\right) \varepsilon}}{d\theta^*} \left( \left(1 + \max_{|i-j|=1} \frac{d_i}{d_j}\right) \frac{\Lambda e^K}{\theta^*} + \frac{d_1}{4} \right),$$

$$|E_p(t) - \bar{E}_p(t)| \leq \frac{2e^K \varepsilon e^{\left(1 + \max_{|i-j|=1} \frac{d_i}{d_j}\right) \varepsilon}}{W\theta^*} \left( \left(1 + \max_{|i-j|=1} \frac{d_i}{d_j}\right) \frac{\Lambda e^K}{\theta^*} + \frac{d_1}{4} \right).$$

## Выводы

Для модели массового обслуживания с одним сервером удалось получить оценки эргодичности и устойчивости.

## Список литературы

1. Zeifman A., Satin Ya., Morozov E., Nekrasova R., Gorhsenin A. On the ergodicity bounds for a constant retrial rate queueing model // arXiv:1506.01468v1 [math.PR], 2015.
2. Зейфман А. И., Бенинг В. Е., Соколов И. А. Марковские цепи и модели с непрерывным временем. – М.: Элекс-КМ, 2008. – 168 с.
3. Leorato S., Orsingher E., Satin Ya., Shilova G., Zeifman A. Some universal limits for nonhomogeneous birth and death processes // Queueing System, 2006. P. 139-151.

# Динамическое распределение транспортных потоков в сети крупного города\*

Д. С. Мазурин<sup>78</sup>

**Аннотация:** Представлена общая итерационная схема динамического распределения потоков в транспортных сетях, включающая в себя динамическую загрузку сети и последующее перераспределение потоков. Для моделирования динамики транспортных потоков предлагается использование гибридного подхода, подразумевающего различный уровень детализации для различных элементов сети и сочетающего преимущества микро- и мезоскопических моделей.

**Ключевые слова:** транспортные потоки; динамическое распределение; равновесное распределение

## Введение

Одной из наиболее актуальных проблем в развитии крупных городов остается транспортная проблема, и ее решение невозможно без применения математического моделирования. Модели транспортных потоков позволяют осуществить прогноз автомобильных и пассажирских потоков в транспортных сетях, оценить различные варианты развития транспортной инфраструктуры и оптимизировать работу транспортной системы. Более того, в последнее время транспортные модели стали важнейшей частью интеллектуальных транспортных систем (ИТС), в частности систем информирования пользователей и систем адаптивного управления дорожным движением в реальном времени.

Наиболее общепотребительным подходом к решению задачи прогноза транспортных потоков в сети крупного города является четырехшаговая схема [1, 2], включающая в себя:

- оценку объемов прибытия и отправления в каждом районе (*trip generation*);

---

\*Работа выполнена при поддержке РФФИ (проект № 13–01–12030)

<sup>78</sup>ФИЦ ИУ РАН (ИСА РАН); 1991 г.р.; mazurin@isa.ru

- расчет матриц межрайонных корреспонденций (*trip distribution*);
- расщепление корреспонденций по способам передвижений (*modal split*);
- распределение корреспонденций по транспортной сети (*trip assignment*).

Как правило, эта схема подразумевает статический подход к распределению потоков по сети, при котором предполагается, что характеристики системы (как спрос, так и предложение) постоянны на протяжении всего периода моделирования, и существует некоторое стационарное распределение транспортных потоков по сети. В противоположность этому, динамическое моделирование подразумевает временные зависимости транспортного спроса, предложения и результирующей загрузки сети, позволяет явно учитывать пропускные способности элементов сети и моделировать возникновение и последующую динамику заторов.

## 1. Динамическое распределение транспортных потоков

Задачу динамического распределения транспортных потоков можно сформулировать следующим образом. Дана матрица корреспонденций  $Q(t) = \{q_{ij}(t)\}$ , зависящая от времени. Для каждой корреспонденции из района  $i$  в район  $j$  с отправлением в момент времени  $t$  необходимо найти множество используемых путей  $\mathcal{K}_{ij}(t)$  и распределение корреспонденции по путям этого множества:

$$\left\{ q_{ijk}(t) > 0 \mid \sum_{k \in \mathcal{K}_{ij}(t)} q_{ijk}(t) = q_{ij}(t), k \in \mathcal{K}_{ij}(t) \right\}.$$

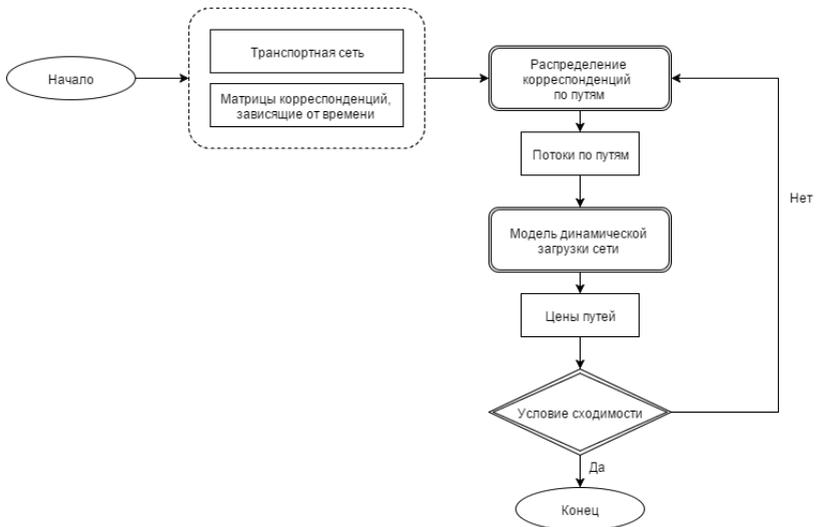
Принципом, который обычно ложится в основу распределения потоков, является принцип пользовательского равновесия Вардропа, согласно которому ни один пользователь не может изменить свой маршрут движения, снизив тем самым цену пути citeSzeto2012. Тогда для каждой пары районов  $i$  и  $j$  и для каждого интервала времени отправления  $t$ :

- все используемые пути из  $i$  в  $j$  имеют одинаковую цену;
- цена любого используемого пути не превосходит цену неиспользуемых путей.

## 2. Общая схема

На Рис. 1 изображена общая схема модели динамического распределения потоков в транспортной сети. Она включает в себя два фундаментальных компонента:

- 1) модель распределения потоков по сети, в которой всем водителям назначается определенный маршрут движения;
- 2) модель динамической загрузки сети, в которой транспортные потоки перемещаются по сети в соответствии с назначенными маршрутами движения.



**Рис. 1.** Общая схема динамического распределения транспортных потоков

Для поиска равновесного динамического распределения транспортных потоков используется итерационный алгоритм. На

каждой итерации модель распределения потоков по сети определяет пути, по которым следуют транспортные средства и их загрузку, то есть частоту отправления транспортных средств по каждому из путей. Имитационная модель динамической загрузки сети воспроизводит динамику транспортных потоков в системе при данной загрузке маршрутов и определяет временные характеристики элементов сети: плотности, скорости движения, задержки в узлах сети и цены всех путей. На основе этих данных происходит следующая итерация перераспределения потоков по сети и т.д.

Исходными данными для модели является транспортная сеть со всеми схемами регулирования и управления движением и зависящие от времени матрицы корреспонденций, которые для каждого момента времени определяют частоту отправления в этот момент из каждого района отправления в каждый район прибытия. Характеристики элементов транспортной сети могут также зависеть от времени ввиду множества факторов: перекрытия отдельных улиц, временных ограничений на движение отдельных категорий транспортных средств, наличия реверсивные полос и т.д.

Для расчета нового распределения потоков используется метод последовательных средних значений [4]. На каждой его итерации сначала формируется множество путей для перераспределения путем добавления нового кратчайшего пути, а затем происходит перенос части спроса с неоптимальных на оптимальные пути до достижения приемлемой сходимости к равновесию.

### **3. Модель динамической загрузки сети**

Модель динамической загрузки сети воспроизводит динамику транспортных потоков в сети и отображает зависящую от времени загрузку путей во временные характеристики элементов сети: плотности, скорости движения, задержки в узлах и цены путей. В основе модели динамической загрузки сети могут лежать различные уровни детализации: распространения потока по сети может воспроизводиться на макро-, микро- или мезоуровне.

Важнейшей характеристикой движения в современных условиях является многополосность. Эффекты перестроений оказывают значительное влияние на пропускную способность элементов, что можно увидеть на транспортных развязках типа «клевер-

ный лист». Макроскопические и мезоскопические модели не всегда позволяют корректно учесть конфигурацию транспортной сети и индивидуальное поведение участников движения. Микроскопический подход описывает поведение каждого водителя, однако чрезвычайно требователен к вычислительным ресурсам и не применим к моделированию транспортной системы крупного города. В связи с этим для расчета загрузки сети при заданном распределении потоков по путям предлагается использовать различный уровень детализации на различных участках сети: на протяженных участках равномерного движения, где детальное описание перестроений не является столь важным, можно ограничиться мезоскопическим уровнем описания, в то время как в узлах сети, на пересечениях и вблизи узких мест необходимо использовать микроскопическую модель. Это позволит сохранить приемлемую точность в воспроизведении динамики потока в сети, существенно сократив время вычислений.

В качестве микроскопической модели используется дискретная модель на основе клеточных автоматов [5], в которую введена более точная процедура принятия решений при смене полосы. В отличие от многих существующих моделей вероятность смены полосы не является постоянной, а зависит от условий движения на соседней полосе. Также были введены два новых механизма, повышающие адекватность модели вынужденных перестроений: нарастание «агрессивности» водителя при приближении к цели и проявление водителями взаимной вежливости.

## Выводы

В работе представлена модель динамического распределения транспортных потоков в сети крупного города. В ее основе лежит итерационный процесс, включающий расчет загрузки сети в соответствии с найденным распределением корреспонденций по путям и последующее перераспределение потоков по сети на основе рассчитанной загрузки. Для реалистичного воспроизведения динамики транспортного потока и возникающих в сети заторов предлагается использовать гибридную модель, сочетающую преимущества микро- и мезоскопических моделей и позволяющую сократить время вычислений. Модель может использоваться в транспортном пла-

нировании для оценки влияния тех или иных мероприятий на работу транспортной системы, а также являться основной составляющей интеллектуальной транспортной системы – в режиме реального времени рассчитывать краткосрочный прогноз загрузки сети и времен в нескольких сценариях, вырабатывать стратегии управления светофорами, реверсивными полосами и другими элементами сети, а также давать рекомендации пользователям по выбору тех или иных маршрутов движения.

## Список литературы

1. *Швецов В. И.* Математическое моделирование транспортных потоков // Автоматика и телемеханика, 2003. No.11. С. 3–46.
2. *Алиев А. С., Мазурин Д. С., Максимова Д. А., Швецов В. И.* Структура комплексной модели транспортной системы г. Москвы // Труды ИСА РАН, 2003. Т. 65. No.1. С. 3–15.
3. *Szeto W. Y., Wong S. C.* Dynamic traffic assignment: model classifications and recent advances in travel choice principles // Central European Journal of Engineering, 2012. Vol. 2, No.1. P. 1–18.
4. *Florian M., Mahut M., Tremblay N.* Application of a simulation-based dynamic traffic assignment model // European Journal of Operational Research, 2008. Vol. 189, No.5. P. 1381–1392.
5. *Мазурин Д. С.* Математическая модель динамики транспортного потока на многополосной магистрали // Автоматика и телемеханика, 2013. No.5. С. 156–166.

# О стационарных характеристиках системы $Geo/Geo/1/r$ с полным обновлением, зависящим от состояния очереди\*

Л. А. Мейханаджян<sup>79</sup>, И. С. Зарядов<sup>80</sup>, Р. В. Разумчик<sup>81</sup>

**Аннотация:** Рассматривается система  $Geo/Geo/1/r$  с полным обновлением, в которую поступает геометрический поток заявок. Имеется накопитель неограниченной емкости и один обслуживающий прибор. Время обслуживания каждой заявки распределено по геометрическому закону. Каждая заявка в момент окончания обслуживания с вероятностью  $q(k)$  опустошает очередь, при условии, что в очереди находится  $k$  заявок. Разработаны метод нахождения основных стационарных показателей функционирования в случаях, когда  $r < \infty$  и  $r = \infty$ , а также с дообслуживанием заявок, опустошающих очередь и без.

**Ключевые слова:** система массового обслуживания; обновление; дискретное время

## Введение

Согласно [1] системы массового обслуживания (СМО) в дискретном времени являются адекватными моделями цифровых систем передачи данных, поскольку позволяют адекватно описывать процессы образования очередей в этих системах, предоставляя возможность учитывать дискретный характер передаваемой информации. В СМО в дискретном времени время считается разделенным на такты фиксированной длины  $h$ , которая может зависеть от специфики рассматриваемой системы. К настоящему времени имеется обширная литература по аналитическим методам анализа

---

\*Работа выполнена при поддержке РФФИ (проекты 13-07-00223 и 15-07-03007).

<sup>79</sup>РУДН; 1990 г.р.; lameykhanaadzhyan@gmail.com

<sup>80</sup>РУДН; 1981 г.р.; к.ф.-м.н.; izaryadov@gmail.com

<sup>81</sup>ФИЦ ИУ РАН (ИПИ РАН); 1984 г.р.; к.ф.-м.н.; rrazumchik@ipiran.ru

СМО в дискретном времени и их применению при анализе конкретных информационно-телекоммуникационных систем [2, 3, 4]. Системы с обновлением, по-видимому впервые, были рассмотрены в работе [5] и с тех пор были предметом активных исследований. В общем случае (обобщенное) обновление представляет моделирование такого эффекта, при котором обслужившаяся заявка, покидая систему, “вытесняет” из системы еще и некоторое число других заявок, ожидающих в очереди. С наступлением каждого события вытеснения заявок связывается некоторая вероятность, которая и называется вероятностью обновления. Согласно [6] обобщенное обновление можно рассматривать как вариант развития идеи отрицательных заявок ([7] и [8]) и катастроф ([9, 10, 11] и [12, 13]). Главное отличие систем с обновлением состоит в том, что потеря заявок из накопителя происходит не в момент поступления другой заявки в систему, а в момент окончания обслуживания на приборе. Кроме того стоит отметить, что обновление может рассматриваться, как вариант механизма активного управления очередью: управляя вероятностями обновления можно обеспечивать соблюдение необходимых показателей качества функционирования (заданный уровень средней длины очереди и потерь). В большинстве работ по системам с обновлением рассматривался случай непрерывного времени. Учитывая то внимание, которое уделяется системам в дискретном времени (см. [1]), представляет интерес рассмотрение СМО с обновлением в дискретном времени, а также выяснение границ применимости для их анализа уже известных (для непрерывного случая) аналитических методов.

## 1. Описание системы

Рассматриваемый вариант системы  $Geo/Geo/1/r$  с полным обновлением представляет собой функционирующую в дискретном времени СМО с одним прибором и накопителем бесконечной емкости для ожидающих заявок. Входящий в систему поток является геометрическим, что означает, что на каждом такте с вероятностью  $a$ , не зависящей от моментов поступления всех предыдущих заявок, поступает новая заявка и начинает обслуживаться, если прибор свободен, становится в очередь, если есть свободные места в очереди, и покидает систему в противном случае. Время обслу-

живания каждой заявки не зависит от всего остального процесса функционирования системы и также распределено по геометрическому закону с параметром  $b$ . Каждая заявка в момент окончания обслуживания с вероятностью  $q(k)$  опустошает очередь (т.е. “вытесняет” из накопителя все заявки), если в системе (не считая обслуженной) находится  $k \geq 0$  заявок. Согласно [14, 6]  $q(k)$  называются вероятностями (полного) обновления. Без ограничения общности, длина такта  $h$  рассматривается равной 1, и в момент окончания такта принята следующая последовательность событий: сначала систему покидает заявка на приборе, если закончилось ее обслуживание, затем, если закончилось обслуживание, происходит обновление с вероятностью  $q(k)$ , зависящей от числа заявок в системе, и потом уже в систему принимается поступившая заявка, если такая имеется. Заявки на обслуживание выбираются из очереди в порядке поступления в систему.

## 2. Основные результаты

Функционирование рассмотренной СМО по времени (по моментам окончания тактов) описывается цепью Маркова со счетным множеством состояний, образованной общим числом заявок в системе в последовательные моменты времени непосредственно после окончания тактов. Были разработаны методы, позволяющие рассчитывать основные стационарные характеристики очереди для случаев, когда  $r < \infty$  и  $r = \infty$ , а также с учетом дообслуживания опустошающей очередь заявки и без. Были найдены аналитические выражения для стационарных распределений числа заявок по времени и по моментам поступления, стационарные распределения времен ожидания и пребывания заявок в системе.

**Таблица 1.** Значения основных показателей производительности системы при  $a = 0.6$ ,  $b = 0.1$ ,  $r = 20$  и различных значениях  $q(k) = q$

	$q = 0.5$	$q = 0.05$	$q = 0.005$
$\pi$	0.13	0.68	0.81
$N$	9.41	18.91	20.70
$N^*$	8.89	18.72	20.59
$v$	15.07	26.92	33.00

В таблице 1 представлены некоторые результаты численных расчетов значений вероятности блокировки ( $\pi$ ), среднего числа заявок в системе по времени ( $N$ ), среднего числа заявок в системе по моментам поступления ( $N^*$ ); среднего времени пребывания заявки в системе ( $v$ ).

## Список литературы

1. *Bruneel H., Fiems D., Walraevens J., Wittevrongel S.* Queueing models for the analysis of communication systems // TOP, 2014. Vol.22. Iss.2. P. 421–448.
2. *Morozov E., Fiems D., Bruneel H.* Stability analysis of multiserver discrete-time queueing systems with renewal type server interruptions // Performance Evaluation, 2011. Vol.68. P.1261–1275.
3. *Wang J., Huang Y., Dai Z.* A discrete-time on-off source queueing system with negative customers // Computers & Industrial Engineering, 2011. Vol.61. P. 1226–1232.
4. *Walraevens J., Steyaert B., Bruneel H.* Analysis of a discrete-time preemptive resume priority buffer // European Journal of Operational Research, 2008. Vol.186. P. 182–201.
5. *Kreinin A.* Queueing Systems with Renovation //Journal of Applied Mathematics and Stochastic Analysis, 1997. Vol.10. Iss.4. P. 431–443.
6. *Зарядов И. С.* Система массового обслуживания GI/M/n/∞ с обобщенным обновлением // Автоматика и телемеханика, 2010. No.4. С. 130–139
7. *Gelenbe E., Glynn P., Sigman K.* Queues with Negative Arrivals // J. Appl. Probab., 1991. Vol.28. P. 245–250.
8. *Бочаров П. П., Д’Анчиче Ч., Мандзо Р., Печинкин А. В.* Анализ многолинейной марковской системы массового обслуживания с неограниченным накопителем и отрицательными заявками // Автоматика и телемеханика, 2007. No.1. С. 93–104.

9. *Dudin A. N., Karolik A. V.* BMAP/SM/1 Queue with Markovian Input of Disasters and Non-Instantaneous Recovery // Performance Evaluat., 2001. Vol.45. Iss.1. P. 19–32.
10. *Дудин А. Н., Кум Ч. С., Семенова О. В.* Оптимальное многопороговое управление входным потоком для системы обслуживания GI/PH/1 с BMAP-потоком отрицательных запросов // Автоматика и телемеханика, 2004. No.9. С. 71–84.
11. *Semenova O. V., Dudin A. N.* M/M/N Queueing System with Controlled Service Mode and Disaster // Automatic Control and Computer Sciences, 2007. Vol.41. Iss.6. P. 350—357.
12. *Semenova O. V.* Optimal Control for a BMAP/SM/1 Queue with MAP-input of Disasters and Two Operation Modes // RAIRO Oper. Res. Recherche Operationnelle, 2004. Vol.38. Iss.2. P. 153–171.
13. *Krishna Kumar B., Krishnamoorthy A., Pavai Madheswari S., Sadiq Basha S.* Transient Analysis of a Single Server Queue with Catastrophes, Failures and Repairs // Queueing Systems: Theory and Applications, 2007. Vol.56. Iss.3. P. 133–141.
14. *Зарядов И. С., Печинкин А. В.* Стационарные временные характеристики системы GI/M/n/∞ с некоторыми вариантами дисциплины обобщенного обновления // Автоматика и телемеханика, 2009, No.12. С. 161–174.

# Имитационная модель сценария совместного использования частот аэропортом и оператором сети 3GPP LTE

А. А. Пономаренко-Тимофеев<sup>82</sup>

**Аннотация:** Для анализа сетей связи часто прибегают к использованию программ имитационного моделирования. В большинстве случаев применение таких программ, как Network Simulator 3, OPNET Riverbed, позволяет моделировать работу как беспроводных, так и проводных сетей. Однако в некоторых случаях функционал данных программ не позволяет провести эксперименты без усложнения сценария, что требует создания дополнительных условий и объектов в модели. В работе рассматривается одна из таких ситуаций на примере моделирования сети LTE (Long Term Evolution) с применением системы совместного использования частот LSA (Licensed Shared Access) в условиях близости к аэропорту.

**Ключевые слова:** имитационное моделирование; LTE; LSA

## 1. Программный комплекс WINTERSim

Для моделирования сценария используется программный комплекс, разработанный коллективом научно-исследовательской группы W.I.N.T.E.R. (Wireless Intelligence for Networking Technology by Engineering and Research) Технологического университета Тампере (Финляндия). Программа позволяет проводить моделирование как проводных, так и беспроводных сетей (сотовые, Wi-Fi сети), создавать сценарии с разными технологиями, объектами, а также собирать статистические данные интересующих характеристик. Возможна работа с системами массового обслуживания и сетями массового обслуживания. Работа с моделью происходит через специальный файл 'scenario', в котором описываются ключевые алгоритмы и объекты, используемые в модели.

---

<sup>82</sup>РУДН; 1992 г.р.; alexeyponomarenko92@gmail.com

Для определения параметров модели (ширина используемой полосы, количество подканалов и уровень теплового шума, размер шага моделирования) используются дополнительные файлы с параметрами. Дополнительные модули написаны на языке Python, кроме физического уровня, реализованного на языке C++. После завершения работы программа выводит результаты в формате, совместимом с Matlab (или другой программой, способной обрабатывать файлы в формате Matlab), что позволяет в дальнейшем произвести более подробный анализ поведения системы.

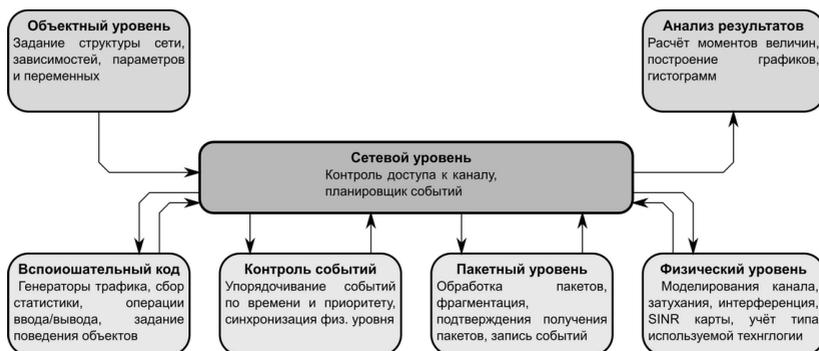
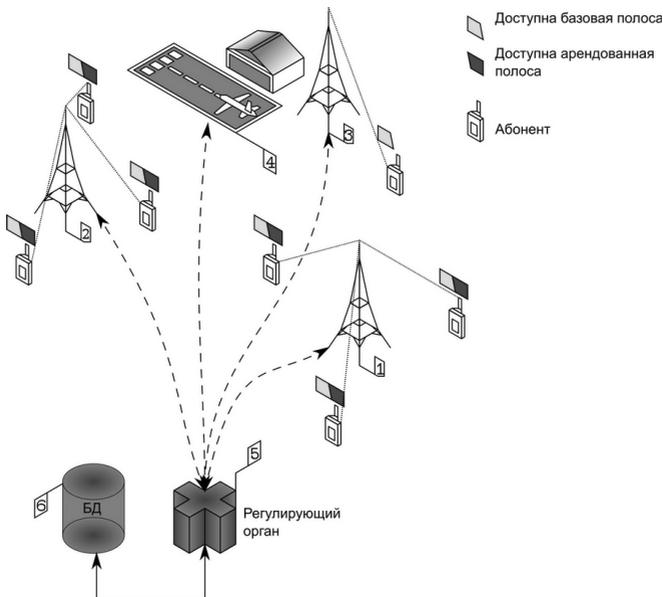


Рис. 1. Схема модулей программного комплекса WINTERSim

## 2. Описание сценария

Совместное использование частот LSA позволяет оператору воспользоваться полосой частот, изначально принадлежащей другой стороне (например, аэродромы, радары, системы связи на железной дороге и другие системы). Владелец имеет приоритет над арендатором-оператором, т.е. в случае, когда владельцу потребуются сданная в аренду полоса, арендатор обязан освободить эту полосу и предоставить ее владельцу. Другой вариант – арендатор может не освобождать полосу, а понизить мощность передаваемого сигнала и отправить абонентам сервисное сообщение с требованием понизить мощность передачи [1]. Данный подход позволяет продолжить работу с абонентами на арендованных полосах и не нарушать условия договора с владельцем. На Рис. 2 изображен один из приме-

ров применения LSA. Оператору принадлежат станции 1,2,3, часть из них использует частоты, изначально принадлежащие аэропорту. Данные о том, какие именно полосы и на какой территории, хранятся в базе данных, находящейся под контролем регулирующего органа. При взлете самолета одна из сот (3) начинает интерферировать с оборудованием самолета, по этой причине арендованные полосы недоступны в данной соте.

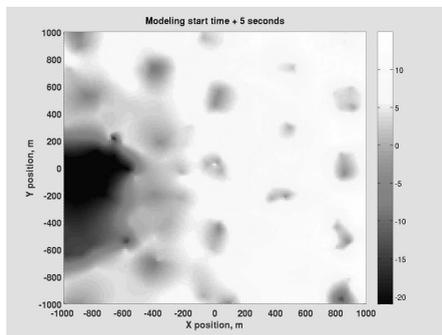


**Рис. 2.** Сценарий совместного использования частот аэропортом и оператором

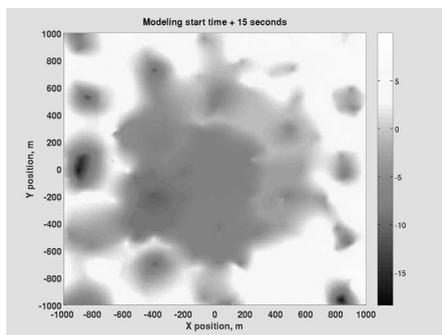
### 3. Анализ характеристик

Рассмотрены различные политики работы с арендованной полосой и проведен их сравнительный анализ по критерию влияния на интерференцию с самолетом.

Пусть сеть состоит из 25 сот (5x5), в каждой из которых есть только LSA частоты. Каждая из сот является шестиграни-

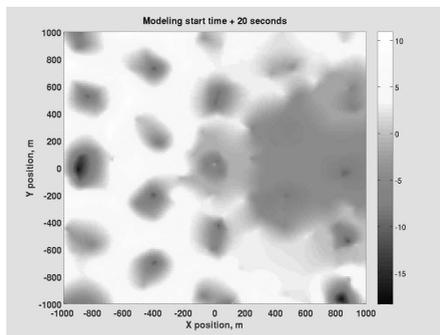


**Рис. 3.** Появление самолета над сетью



**Рис. 4.** Радиотень от самолета

ком, расстояние между центрами сот составляет 500 метров. В каждой соте обслуживается 30 абонентов, генерирующих восходящий трафик. Исходным владельцем полосы является аэропорт, использующий арендуемую полосу для связи с самолетом (телеметрические системы для сопровождения самолета). Самолет набирает высоту по мере пролета над сетью и проводит измерения мощности принимаемого сигнала от сети. Одна из политик работы с арендованной полосой такая, что при превышении интерференционного порога на приемнике самолета, оператор понижает мощность передачи на базовых станциях и мощность передатчика абонента. Данный подход позволяет продолжать обслуживать абонентов, находящихся в непосредственной близости к базовой станции. На



**Рис. 5.** Выход самолета за границы сети

Рис. 3-5 приведены несколько карт мощностей передачи абонентских устройств для рассматриваемого сценария.

### Благодарности

*Автор выражает благодарность А.В. Пяттаеву (Технологический университет Тампере, Финляндия) и М.А. Герасименко (Технологический университет Тампере, Финляндия) за рекомендации по работе с комплексом программных средств WINTERSim, а также И.А. Гудковой (Российский университет дружбы народов) и С.Д. Андрееву (Технологический университет Тампере, Финляндия) за ценные советы при подготовке работы.*

### Список литературы

1. Mrs. Sonia, Malik N., Rekhi P.K.S., Malik S.S. Uplink Power Control Schemes in Long Term Evolution // IJEAT 2014. P. 260–264.
2. Borodakiy V. Y., Samouylov K. E., Gudkova I. A., Ostrikova D. Y., Ponomarenko A. A., Turlikov A. M., and Andreev S. D. Modeling unreliable LSA operation in 3GPP LTE cellular networks // Proc. of the 6th International Congress on Ultra Modern Telecommunications and Control Systems ICUMT-2014 (October 6–8, 2014, St. Petersburg, Russia). – IEEE. – 2014. – P. 390–396.

Научное издание

ТРУДЫ ВТОРОЙ МОЛОДЕЖНОЙ  
НАУЧНОЙ КОНФЕРЕНЦИИ  
«ЗАДАЧИ СОВРЕМЕННОЙ ИНФОРМАТИКИ»

Под редакцией к.ф.-м.н. Горшенина А. К.

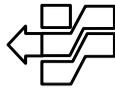
Оригинал-макет подготовлен в ФИЦ ИУ РАН

Подписано в печать 26.10.2015

Тираж 120 экз.

Заказ №15-05

Издано ФИЦ ИУ РАН  
119333, Москва, ул. Вавилова, д.44, корп.2



ISBN 978-5-91993-054-9