

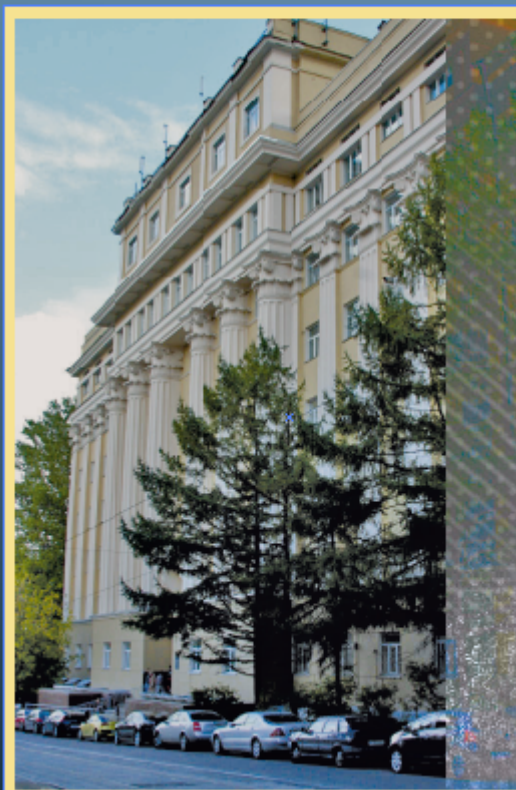
ИНФОРМАЦИОННО- ТЕЛЕКОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ И МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ВЫСОКОТЕХНОЛОГИЧНЫХ СИСТЕМ

МАТЕРИАЛЫ
Всероссийской
конференции
с международным
участием

Москва, РУДН
16–20 апреля
2018 г.



Москва
Российский университет
дружбы народов
2018



**ИНФОРМАЦИОННО-
ТЕЛЕКОММУНИКАЦИОННЫЕ
ТЕХНОЛОГИИ И МАТЕМАТИЧЕСКОЕ
МОДЕЛИРОВАНИЕ
ВЫСОКОТЕХНОЛОГИЧНЫХ СИСТЕМ**

**Материалы
Всероссийской конференции
с международным участием**

Москва, РУДН, 16–20 апреля 2018 г.

УДК 004:007(063)
ББК 32.81
И74

Конференция проводится в рамках реализации Программы повышения конкурентоспособности РУДН «5-100», проект М 2.4.1.П1 «Организация и проведение НТМ, повышающих международный и всероссийский уровень признания ученых РУДН».

Организатор конференции: Российский университет дружбы народов
Соорганизаторы конференции:

Московский технический университет связи и информатики (МТУСИ);
Институт проблем управления им. В. А. Трапезникова Российской академии наук (ИПУ РАН);
Лаборатория информационных технологий Объединенного института ядерных исследований (ЛИТ ОИЯИ);
Федеральный исследовательский центр «Информатика и управление» Российской академии наук (ФИЦ ИУ РАН)

Программный комитет: Самуйлов К. Е., д.т.н., проф., РУДН — председатель программного комитета; Мележик В. С., д.ф.-м.н., ЛТФ ОИЯИ, РУДН — сопредседатель программного комитета; Севастьянов Л. А., д.ф.-м.н., проф., РУДН — сопредседатель программного комитета; Чукарин А. В., к.ф.-м.н., доцент, РУДН — сопредседатель программного комитета; Гудкова И. А., к.ф.-м.н., доцент, РУДН — секретарь программного комитета; Андреев С. Д., к.т.н., РУДН, Tampere University of Technology, г. Тампере, Финляндия; Башарин Г. П., д.т.н., проф., РУДН; Боголюбов А. Н., д.ф.-м.н., проф., МГУ; Виницкий С. И., д.ф.-м.н., проф., ЛТФ ОИЯИ; Вишневецкий В. М., д.т.н., проф., ИПУ РАН; Гайдамака Ю. В., д.ф.-м.н., доцент, РУДН; Гнатич М. М., проф., Pavol Jozef Šafárik University in Košice, Словакия; Гольдштейн Б. С., д.т.н., проф., СПб ГУТ; Горшенин А. К., к.ф.-м.н., доцент, ФИЦ ИУ РАН; Дружинина О. В., д.ф.-м.н., проф., ФИЦ ИУ РАН; Ефимушкин В. А., к.ф.-м.н., доцент, ОАО «Интеллект Телеком»; Кореньков В. В., д.т.н., проф., ЛИТ ОИЯИ; Королькова А. В., к.ф.-м.н., доцент, РУДН; Крынев А. В., д.ф.-м.н., проф., НИЯУ «МИФИ»; Кулябов Д. С., к.ф.-м.н., доцент, РУДН; Кучерявый А. Е., д.т.н., проф., СПб ГУТ; Кучерявый Е. А., к.т.н., проф., НИУ ВШЭ; Ланеев Е. В., д.ф.-м.н., проф., РУДН; Мартикайнен О. Е., проф., Service Innovation Research Institute, г. Хельсинки, Финляндия; Назаров А. А., д.т.н., проф., ТГУ; Наумов В. А., проф., Service Innovation Research Institute, г. Хельсинки, Финляндия; Осипов Г. С., д.ф.-м.н., проф., ФИЦ ИУ РАН; Пузынин И. В., д.ф.-м.н., проф., ЛИТ ОИЯИ; Пшеничников А. П., к.т.н., проф., МТУСИ; Ромашкова О. Н., д.т.н., проф., МГПУ; Севастьянов А. Л., к.ф.-м.н., РУДН; Соченков И. В., к.ф.-м.н., РУДН, ФИЦ ИУ РАН; Степанов С. Н., д.т.н., проф., МТУСИ; Сушенко С. П., д.т.н., проф., ТГУ; Хачумов В. М., д.т.н., проф., ФИЦ ИУ РАН; Цирулев А. Н., д.ф.-м.н., проф., ТвГУ; Цитович И. И., д.ф.-м.н., доцент, ИПШИ РАН; Чудунбаатар О., д.ф.-м.н., ЛИТ ОИЯИ; Шоргин С. Я., д.ф.-м.н., проф., ФИЦ ИУ РАН; Щетинин Е. Ю., д.ф.-м.н., проф., СТАНКИН

Оргкомитет:

Председатель: Васильев С. А., к.ф.-м.н. (РУДН).
Сопредседатели: Диваков Д. В., к.ф.-м.н. (РУДН); Острикова Д. Ю., к.ф.-м.н. (РУДН); Стрельцова О. И., к.ф.-м.н., с.н.с. (ЛИТ ОИЯИ).
Члены оргкомитета: Никитина Е. В., к.х.н., зам. декана по науке (РУДН); Малых М. Д., к.ф.-м.н. (РУДН); Демидова А. В., к.ф.-м.н. (РУДН); Тютюнник А. А. (РУДН)

И74 **Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем** : материалы Всероссийской конференции с международным участием. Москва, РУДН, 16–20 апреля 2018 г. — Москва : РУДН, 2018. — 434 с. : ил.

ISBN 978-5-209-08641-3

Содержание

Теория телеграфика и её применения

Агеев К. А., Сопин Э. С. Разработка средства имитационного моделирования ресурсных систем массового обслуживания	17
Аду К. И., Маркова Е. В. Рекуррентный алгоритм для расчета показателей эффективности СМО с приборами, находящимися под влиянием случайной среды	20
Багаева Н. В. Анализ модели процесса обслуживания клиентов страховой компании с ограниченным сроком действия договоров страхования	23
Богданова Е. В., Зарядов И. С., Милованова Т. А. Математическая модель системы с рекуррентным входящим потоком, повторным обслуживанием и обобщенным обновлением	26
Гольская А. А., Маркова Е. В., Дзантиев И. Л. Модель соты беспроводной М2М сети с равномерно распределенными на плоскости источниками заявок в виде СМО	29
Горшенин А. К. Об исследовании параметров метеорологических моделей на основе паттернов	32
Дудин А. Н., Дудин С. А., Дудина О. С. Модели массового обслуживания для описания работы узлов беспроводных сенсорных сетей с самогенерацией энергии	35
Ибрагимов Б. Г., Гумбатов Р. Т., Ибрагимов Р. Ф. Исследование управления информационными и сетевыми ресурсами в мультисервисных сетях связи	38
Ибрагимов Б. Г., Керимов В. Р., Гаджизаде Н. К., Исаев А. М. Исследование ресурсов в мультисервисных сетях связи на базе единого инфокоммуникационного пространства	44
Карнаухова Е. А., Иванова Д. В., Маркова Е. В. Модель схемы распределения радиоресурсов в сети LTE в виде СМО с ненадежными приборами и конечной очередью	49
Крупко О. С., Москалева Ф. А., Ковальчуков Р. Н., Самуйлов А. К., Молчанов Д. А. Математическая модель звена доступа в сетях связи миллиметрового диапазона для одноадресных и многоадресных соединений	52
Назаров А. А., Анисимова А. А. Асимптотический анализ второго порядка двухфазной гибридной СМО	54
Никитина В. В., Хуракай Д. М. Об этапах интеллектуального анализа данных при построении и совершенствовании бизнес-процессов (Process Mining)	57
Орбелян Т. С., Зарядов И. С., Милованова Т. А. Система с выбором кратчайшей очереди при поступлении заявок и максимальной очереди при обслуживании заявок	60
Погосян Ш. Г. Анализ модели процесса обработки разнотипных исков в страховой компании	63
Ромашкова О. Н., Пономарева Л. А., Василюк И. П. Применение информационных технологий для анализа показателей рейтинговой оценки вуза	65
Семёнова О. В., Буй Д. Т. Адаптивные дисциплины опроса в системах поллинга и их имитационное моделирование	69
Синицын И. Э. Анализ вероятности успешного установления соединения по радиоканалу случайного доступа с использованием виртуальных преамбул	72
Уанкпо Г. Ж. К., Козырев Д. В. Модель надёжности однородной системы передачи данных облегчённого резервирования с произвольным распределением времени ремонта элементов	75
Халина В. А. Анализ временных показателей бизнес-процессов	78
Цурлуков В. В., Зарядов И. С., Милованова Т. А. К анализу системы массового обслуживания с ресурсами, функционирующей в случайном окружении	81
Сети связи следующего поколения: управление, качество, архитектура	
Абаев П. О., Царев А. С. Модель для анализа показателей эффективности ядра сети 5G с виртуализацией и гистерезисным управлением	87

Андреев С. Д. Прямое взаимодействие пользователей в мобильных социальных сетях	90
Бахтин А. А. Разработка протокола передачи данных в сетях MANET	93
Богданова Е. Ю. Численный анализ простейшей модели транспортного потока в умном городе	96
Волков А. С., Солодков А. В., Баскаков А. Е., Муратчаев С. С. Разработка двухранговой помехоустойчивой системы радиосвязи на основе кодового разделения каналов	98
Волков А. С., Лазарев А. В. Исследование и выбор помехоустойчивых кодов в соответствии с требованиями сетей мобильной связи пятого поколения	101
Волынец А. С., Гарибян А. А. Реализация модели случайного блуждания для описания движения объектов в ограниченном пространстве	104
Гайдамака А. А., Чухно Н. В., Чухно О. В. Алгоритм принятия решений с помощью вычислений со словами для задач большой размерности	107
Гольдштейн В. С., Соколов Н. А. Задачи безопасности нормализуемого QoS в сетях следующего поколения	110
Гудкова И. А., Романовская Ю. А. Оценка производительности работы LSA на примере модели отдельно взятой соты	115
Дараселия А. В. Анализ механизмов повышения энергоэффективности облачных систем	118
Жданков А. Н., Гайдамака Ю. В. Об одном алгоритме моделирования движения абонентов в беспроводной сети	121
Карачанская Е. В., Соседова Н. И. Разработка метода выявления аномалий сетевого трафика	124
Кименчежи В. В., Козырев Д. В. Имитационное моделирование мобильности участников соединения в гетерогенной сети беспроводной передачи данных с ограничением на пересечение границ и ненадежными объектами	127
Кутбитдинов С. Ш., Лохмотко В. В., Рудинская С. Р. Гармонизация параметров IMS методом взаимной оптимизации по критериям производительности и потерь	131
Макеева Е. Д., Харин П. А., Поляков Н. А., Маркова Е. В., Гудкова И. А., Галинина О. С. Анализ установления соединения пар передатчик-приемник на mmWave	134
Мачнев Е. А., Бегишев В. О. Имитационное моделирование уличных точек доступа, функционирующие на миллиметровом диапазоне частот	137
Мацкевич И. А. Исследование наработки сенсорной сети полиграфического комбината до отказа	140
Половов М. П., Бесчастный В. А., Острикова Д. Ю., Гудкова И. А. Численный анализ оптимальной скорости передачи данных в сети с технологией мультимедиа методом перебора	143
Разгоянев В. А., Мокров Е. В., Самуйлов К. Е. Математическая модель среднего времени передачи данных подвижным пользователям в сети LTE	146
Ромашкова О. Н., Самойлов В. Е. Определение дальности гарантируемой радиосвязи в беспроводных телекоммуникационных сетях стандарта IEEE 802.11 с использованием программы ring	150
Савич В. Н., Бесчастный В. А., Острикова Д. Ю., Гудкова И. А. Модель мобильности в виде СЕМО с «пассивным» узлом	154

Прикладные информационные системы и технологии

Волков С. С. Система распознавания рукописных букв на основе нейронной сети	159
Григорьева Т. В., Жуков В. В. Перспективы применения кластерного анализа для реинжиниринга клинических бизнес-процессов	162
Жуков В. В., Аронов Д. А., Семушина С. Г., Моисеева Е. В. Применение машины опорных векторов для предсказания возникновения рака молочной железы в мышинной модели	165
Зубрихина М. О., Молодченков А. И. Разработка метода поддержки принятия решений о выборе тактики хирургического лечения грыжи диска	168
Катрич А. В. Пример применения OLAP	171

Козловский А. Н., Жуков В. В. Вычисление апостериорных вероятностей в психопатологии на основе метаанализа медицинских публикаций	174
Коняев Е. Ю. Автоматизация оценивания вопросов открытого типа	177
Крескин А. Д. Выявление источников заимствования для документа с использованием моделей дистрибутивной семантики	180
Кузнецов Е. А., Фомин М. Б. Проектирование многомерных информационных систем с использованием методологии «Data Vault»	183
Лукин А. В. Построение системы проведения мета-анализа медицинских исследований для получения ответа на клинический вопрос (на примере печеночной недостаточности)	186
Матвеева Е. А. Методы синтеза текстов на естественном языке в заданном стиле	189
Молодченков А. И. Модуль базы знаний в системе управления здоровьем людей	192
Новикова Д. С. Разрешение неоднозначности авторства публикаций на основе коллекции документов поисковой системы Eхastus Expert	195
Павлюкова А. В. Использование статистических характеристик сигнала в построении аудиоидентификатора музыкальных произведений для поиска нечетких соответствий	198
Пальчевский А. И., Молодченков А. И. Разработка методов моделирования динамики изменения ишемической болезни мозга путем применения алгоритмов 3d моделирования и морфинга изображений	201
Паршина К. С. Алгоритмы удаления невидимых линий при построении изображений трехмерных тел	204
Перетягько О. Д., Молодченков А. И. Оценка соответствия экземпляра процесса его обобщенной схеме	207
Салпагаров С. И., Гончаров Л. М., Мардашев А. М. Векторная задача о сочетаниях на гиперграфах	210
Салпагаров С. И., Исаев Ю. Д. Оптимизационная модель распределения потоков данных R2P-телевидения на гиперграфах	213
Сингх Л., Молодченков А. И. Разработка метода оценки близости пептидов	216
Скрынник К. В. Исследование метода классификации новостей на основе дистрибутивной семантики	219
Тханг Фам К., Копылов А., Тхао Чан Т. Т. Вычислительно эффективный алгоритм восстановления изображений с использованием полной вариации	222
Хачумов М. В., Хачумов В. М. Задача безопасного формирования заданного строя БПЛА (плоский случай)	226
Храбров Р. Н., Фомин М. Б. Визуализация потока жидкости через тонкостенный сосуд с переменным сечением	229
Чистова Е. В. Моделирование естественно-языковой коммуникации с помощью глубокого обучения	232
Ядринцев В. В., Соченков И. В. Полнотекстовый классификатор патентных документов	235
Янко Ю. Д. Проектирование базы данных для системы автоматизации процесса голосования	238

Распределённые, высокопроизводительные вычисления и аналитика больших данных

Адам Г., Беляков Д. В., Валя М., Гончаров П. В., Зрелов П. В., Кореньков В. В., Матвеев М. А., Подгайный Д. В., Стрельцова О. И. Виртуальные рабочие столы гетерогенного кластера HybridIT	245
Башапин М. В., Земляная Е. В., Лукьянов К. В. Параллельная реализация модели микроскопического оптического ядро-ядерного потенциала на основе технологий MPI и OPENMP	248
Добрынин В. Н., Миловидова А. А., Соколов И. А. Оценка адекватности модели и объекта исследования	252
Дубнов Ю. А. Перекрестная энтропия для отбора признаков в задачах анализа данных	257

Зуев М. И., Матвеев М. А., Подгайный Д. В., Стрельцова О. И., Торосян Ш. Г. HLLT-VDI – новый сервис IT-экосистемы гетерогенного кластера HybriLIT для работы с прикладным программным обеспечением	260
Мажитова Е. М., Балашов Н. А., Баранов А. В., Кутовский Н. А., Семенов Р. Н. Интеграция облачных инфраструктур ЛИТ ОИЯИ и Астанинского филиала ИЯФ	262
Никольский И. М. Суперкомпьютерное моделирование функционирования беспроводной сенсорной сети в ситуации потери связи с базовой станцией	266
Рахмонова А. Р. Анализ производительности компьютерного моделирования физических процессов в системе длинных джозефсоновских переходов на кластере HybriLIT	269
Чеботов А. И., Авзалов Р. Л., Чернышев Я. М., Герценбергер К. В., Бутенко Ю. А. Разработка сервиса VM@N-Webapp для обработки и визуализации информации об эксперименте на фиксированной мишени VM@N	273

Математическое моделирование

Аль-Натор М. С., Аль-Натор С. В. Оптимальная самофинансируемая много-периодная портфельная стратегия с учетом комиссии	279
Амирханов И. В., Саркар Н. Р. Об одном методе приближенного решения задачи трех тел	282
Апреутесей А.-М. Ю. Простая модель системы с активным управлением очередью по алгоритму типа RED	285
Ахметханова Д. М., Зарядов И. С. Вероятностные и статистические методы, применяемые для оптимизации работы парковочных систем и связанные с распознаванием номерных знаков	288
Байло Н. В., Матюшенко С. И., Пяткина Д. А. Статистический анализ инфляции в современной России	291
Боровских Д. А., Дилле Гбавле Люк-Дональд, Петров В. А., Савин А. С., Садовников В. К., Хохлов А. А. Применение метода SSA к различным прикладным задачам	294
Бочкарев В. В., Крынев А. В., Смирнов Д. С., Шарафутдинов Р. Б. Информационно-математическая модель поддержки принятия решений по выводу из эксплуатации объектов использования атомной энергии	297
Будочкина С. А. О скобках Пуассона в пространствах V_n -потенциалов	300
Буурулдай А. Э. О функции волатильности в модифицированной модели дисперсии с постоянной эластичностью (CEV)	302
Батгэрэл Б., Димова С. Н., Пузынин И. В., Пузынина Т. П., Тухлиев З. К., Христов И. Г., Христова Р. Д., Шарипов З. А. Моделирование дальнего действия тепловых изменений металла, облучаемого нанокластерами меди	305
Велиева Т. Р., Завозина А. В., Королькова А. В. Определение коэффициентов гармонической линеаризации для детерминистической нелинейной системы с управлением	308
Велиева Т. Р., Завозина А. В., Королькова А. В. NS-2 как средство верификации амплитудно-частотных характеристик системы с управлением	311
Геворкян М. Н., Демидова А. В., Демидова Т. С., Соболев А. А. Моделирование распространения сетевых червей с помощью стохастических дифференциальных уравнений	314
Геворкян М. Н., Демидова А. В., Королькова А. В., Кулябов Д. С., Гостев И. М. Проблемы реализации стохастических численных методов Рунге–Кутты	318
Гоголева С. Ю., Зайнетдинова Л. Г. Об одном итерационном методе для решения плохо обусловленных систем линейных алгебраических уравнений с разреженными матрицами	321
Грачев Д. А., Емельянова Е. Ю., Попова Е. В., Хохлов А. А. Разработка модели биоимпедансометра	324
Дилле Гбавле Люк-Дональд, Емельянова Е. Ю., Попова Е. В., Хохлов А. А. Обработка данных акселерометров мобильных устройств для классификации движений	326

Демидова А. В., Дружинина О. В., Масина О. Н., Мияйлович Н., Ячимович М. Синтез и анализ многомерных математических моделей популяционной динамики	329
Денисенко А. П., Мингажитдинова Э. Ф., Петров В. А., Савин А. С., Хохлов А. А. Оптимизация стоимости транзакций в блокчейне Ethereum	333
Диваков Д. В., Севастьянов А. Л., Белов М. П. Постановка задачи расчета направляемых мод градиентного в поперечном направлении волновода	336
Древицкий А. С., Диваков Д. В., Котоков М. М. Задача символично-численного расчета собственных значений вытекающих мод в регулярном одно-родном открытом волноводе	339
Егоров А. А. Исследование свойств жидкокристаллических волноводных структур	342
Зотова П. А., Жукова Л. В., Пяткина Д. А. Анализ востребованных форматов заданий для онлайн-курсов среди студентов разных учебных заведений	345
Касимов Ю. Ф., Тимербаев М. И. Оптимальные веса рискованных активов в индексных портфелях постоянной пропорции	348
Коновалов М. Г., Разумчик Р. В. Управление марковской цепью с непрерывным ограниченным множеством состояний, оптимизирующее стационарное распределение цепи	351
Крянев А. В., Слива Д. Е., Матохин В. В. Характеристики неравномерности распределения ресурсов в экономических системах	355
Кузев Я. Ю. Комплекс программ для решения нелинейных дифференциальных уравнений в частных производных первого порядка	358
Кулябов Д. С., Геворкян М. Н., Королькова А. В. Численный анализ геометризованных уравнений Максвелла	361
Лесик П. А., Романов А. А., Чистотина Д. А. Об одной обратной задаче для метагармонического уравнения	364
Лижненко Н. М., Морозова Д. А. Задача продолжения ньютоновского потенциала в цилиндре кругового сечения	366
Мысина Ю. А. Повышение пенсионного возраста и решение проблемы снижения численности занятого населения	368
Пермякова Ю. С. Рыночные риски коммерческого банка: методы оценки и управления	371
Попков А. Ю. Метод генерации случайных векторов в заданными вероятностными характеристиками	374
Пылаева А. Н. Моделирование ценообразования на рынке смартфонов	377
Пяткина Д. А., Матюшенко С. И., Зотова П. А., Казандаева А. П. Анализ цен на ноутбуки с помощью эконометрических моделей	381
Рассахан Н. Д., Щетинин Е. Ю. Оценка зависимостей тяжелых хвостов при помощи аппарата копул для анализа осадков	384
Севастьянов Л. А., Васильева Д. Г. Стохастическая динамика роста популяций	387
Талагаев Ю. В. Оценка состояния нелинейных сетевых систем управления, представленных нечеткими моделями Такаги-Сугено	390
Тевелева Е. А., Третьяков Н. П., Терлецкий А. Я. Применение фазового анализа временных рядов для выделения макроэкономических циклов на примере динамики курсов валют	393
Третьяков Н. П., Кочанов А. А. Эффективное уравнение Шрёдингера для одномерной системы с быстро осциллирующими граничными условиями	396
Шебанова К. В., Хамидулина К. И. Устойчивое вычисление нормали к поверхности, заданной приближенно	399
Чупритский В. К., Диваков Д. В. Моделирование прохождения световой волны через открытый планарный волновод методом «ящика Дирихле»	401
Шорохов С. Г., Буурулдай А. Э. О модели локальной волатильности с гиперболическим синусом	404
Шуниц Е. В. Сравнение Online-Q и Experience Replay для обучения нейронной сети в мультиагентной среде	407
Щетинин Е. Ю. О некоторых свойствах непараметрических оценок коэффициентов хвостовой зависимости	410

Щегинин Е. Ю., Бережков М. В., Любин П. Г. Разработка эффективного алгоритма краткосрочного прогнозирования электропотребления с использованием метода ансамбля	413
Юй Ин Методика отыскания алгебраических интегралов дифференциальных уравнений первого порядка	416
Бабин А. В. Обзор методов решения задач мониторинга производительности и доступности ИТ структуры предприятия	419
Конькова Е. В. К анализу системы массового обслуживания фазового типа в частном случае	422
Авторский указатель	425

Contents

Teletraffic theory and its applications

Ageev K. A., Sopin E. S. Development of the Simulation Tool for Resource Queueing Systems	17
Adou K. Y., Markova E. V. Recursive Algorithm for Calculating Indicators of Queueing System Model with Devices under Influence of a Random Environment	20
Bagaeva N. V. The Analysis of Models of Insurance Contracts with a Limited Term of Service	23
Bogdanova E. V., Zaryadov I. S., Milovanova T. A. The Mathematical Model of the Retrieval Queueing System with General Renovation and Recurrent Input Flow	26
Golskaia A. A., Markova E. V., Dzantiev I. L. Scheme Model of Wireless M2M Network with Uniformly Distributed Request Sources in the Plane	29
Gorshenin A. K. On Investigation of Parameters of Meteorological Models Based on Patterns	32
Dudin A. N., Dudin S. A., Dudina O. S. Queueing Models for Description of Operation of Nodes of Wireless Sensor Networks with Energy Harvesting	35
Ibrahimov B. G., Humbatov R. T., Ibrahimov R. F. Research of Information Management and Network Resources in Multiservice Communication Networks	38
Ibrahimov B. G., Karimov V. R., Hajizade N. K., Isayev A. M. The Investigation Resources in Multiservice Communication Networks on the Basis of the Single Infocommunication Space	44
Karnauhova E. A., Ivanova D. V., Markova E. V. Resource Allocation Scheme Model in LTE Network as Queueing System with Unreliable Servers and Finite Queue	49
Krupko A. S., Moskaleva F. A., Kovalchukov R. N., Samuilov A. K., Moltchanov D. A. The Mathematical Model of the mmWave Access Point for Two Types of Traffic	52
Nazarov A. A., Anisimova A. A. Second-Order Asymptotic Analysis of a Two-Phase Hybrid Queueing System	54
Nikitina V. V., Khurakay D. M. About the Stages of Data Mining in Creation and Improving Business Processes	57
Orbelyan T. S., Zaryadov I. S., Milovanova T. A. The System with the Selection of the Shortest Queue for Arrival and the Queue with the Maximum of Tasks for Servicing	60
Pogosian Sh. G. The Analysis of the Model of Processing Different Types of Claims in an Insurance Company	63
Romashkova O. N., Ponomareva L. A., Vasilyuk I. P. Application of Information Technology for the Analysis of the Rating of the University	65
Semenova O. V., Bui D. T. Adaptive Polling Disciplines and Their Simulation Models	69
Sinitsyn I. E. Success Access Probability Analysis Using Virtual Preambles Via Random Access Channel	72
Houankpo H. G. K., Kozyrev D. V. Reliability Model of a Homogeneous Warm-Standby Data Transmission System with General Repair Time Distribution	75
Khalina V. A. Business Processes' Time Metrics Analysis	78
Tsurlukov V. V., Zaryadov I. S., Milovanova T. A. To the Analysis of the Queueing System Operating in the Random Environment with Resource Allocation	81

Next-generation networks: management, quality, architecture

Abaev P. O., Tsarev A. S. Design and Performance Evaluation of Virtualized 5G Network Core Model with Implemented Hysteresis Control	87
Andreev S. D. Direct Connectivity of Users in Mobile Social Networks	90
Bakhtin A. A. Development of Data Transfer Protocol in MANET	93
Bogdanova E. Y. Numerical Analysis of the Simplest Model of Traffic Flow in a Smart City	96
Volkov A. S., Solodkov A. V., Baskakov A. E., Muratchaev S. S. Development of a Two-Level Anti-Jam Radio System Based on DSSS	98

Volkov V. S., Lazarev A. V. Research and Selection of Channel Codes in Accordance with the Requirements of Mobile Networks of the Fifth Generation	101
Volynets A. S., Garibyan A. A. A Realisation of Random Waypoint Mobility Model Describing the Objects Movement in a Limited Space	104
Gaidamaka A. A., Chukhno N. V., Chukhno O. V. On Group Decision Making Method for Big Data Processing Environment	107
Goldstein B. S., Sokolov N. A. Tasks of Ensuring Normalized QoS in Next Generation Networks	110
Gudkova I. A., Romanovskaya J. A. Performance Evaluation of Dynamic LSA Operation Through a Model of a Stand-Alone Cell	115
Daraseliya A. V. On the Analysis of Mechanisms for Improving Cloud Computing System Energy Efficiency	118
Zhdankov A. N., Gaidamaka Yu. V. About One Algorithm of Motion Simulations of Subscriber in Wireless Network	121
Karachanskaya E. V., Sosedova N. I. A Construction of a Method for Detecting Anomalies of Network Traffic	124
Kimenezhi V. V., Kozyrev D. V. Mobility Simulation of Connected Objects in a Heterogeneous Wireless Data Transmission Network	127
Kutbitdinov S. Sh., Lokhmotko V. V., Rudinskaya S. R. Harmonization of IMS Parameters by Mutual Optimization by Criteria for Performance and Losses	131
Makeeva E. D., Harin P. A., Polyakov N. A., Markova E. V., Gudkova I. A., Galinina O. S. Analysis of the Connection Setup of Transmitter-Receiver Pairs on mmWave	134
Machnev E. A., Begishev V. O. Simulation Modeling of Street Access Points Operating on the Millimeter Frequency Band	137
Matskevich I. A. Research on Working Time before Failure of the Sensor Network Installed on a Printing Plant	140
Polovov M. P., Beschastnyi V. A., Ostrikova D. Yu., Gudkova I. A. Numerical Analysis of the Optimum Speed of Data Transmission in the Network with Multiplier Technology by Bruteforce	143
Razgonyayev V. A., Mokrov E. V., Samouylov K. E. Mathematical Model of Average Data Transfer Time for Mobile Users in LTE Network	146
Romashkova O. N., Samoylov V. E. Determination of the Range of the Guaranteed Radio Communication in Wireless Telecommunication Networks of IEEE 802.11 Standard with the Use of Ping Program	150
Savich V. N., Beschastnyi V. A., Ostrikova D. Yu., Gudkova I. A. Model of Mobility as a Kind of Queuing Net with "Passive" Unit	154

Applied information systems and technologies

Volkov S. S. The System of Handwritten Letters Recognition Based on Neural Network	159
Grigoryeva T. V., Zhukov V. V. Prospects of Cluster Analysis Application for Reengineering of Clinical Business Processes	162
Zhukov V. V., Aronov D. A., Semushina S. G., Moiseeva E. V. Application of Support Vector Machine to Predict the Onset of Breast Cancer in Mouse Model	165
Zubrikhina M. O., Molodchenkov A. I. Development of the Method of Supporting Decision-Making about the Choice of Tactics of Surgical Treatment of a Herniated Disc	168
Katrlich A. V. The Example of Using OLAP	171
Kozlovskii A. N., Zhukov V. V. Meta-Analysis of Medical Publications on Psychopathology: a Posterior Probability Calculation	174
Konyaev E. Yu. Automated Scoring of Short-Answer Questions	177
Kreskin A. D. Source Retrieval for a Document Using Models of Distributive Semantics	180
Kuznetsov E. A., Fomin M. B. Designing Multidimensional Information Systems Using the Data Vault Methodology	183
Lukin A. V. Development of the System for Conducting a Meta-Analysis of Medical Research to Obtain a Response to a Clinical Question (on the Example of Liver Failure)	186
Matveeva E. A. Approaches to Natural Language Generation with Style	189

Molodchenkov A. I. Knowledge Base Module of the System for Human Health Management	192
Novikova D. S. Author's Name Disambiguation on the Document Collection of Exactus Expert	195
Pavlyukova A. V. Using the Statistical Characteristics of a Signal in the Construction of an Audio Identifier of Musical Works to Search for Fuzzy Matches	198
Palchevskii A. I., Molodchenkov A. I. Development of Methods Model of Dynamics of Change of a Chemical Disease of a Brain by Application of the Algorithm of 3D Model and Moving Images	201
Parshina K. S. Algorithms for Removing Invisible Lines in Constructing Images of Three-Dimensional Objects	204
Peretyatko O. D., Molodchenkov A. I. Conformance Checking of a Process Instance to its General Scheme	207
Salpagarov S. I., Goncharov L. M., Mardashev A. M. Vectorial Problem on Combinations on Hypergraphs	210
Salpagarov S. I., Isaev Yu. D. Multicriteria Model of Distribution of P2P-TV Data Streams on Hypergraphs	213
Singh L., Molodchenkov A. I. Development of Method for Estimimation of Peptide Similarity	216
Skrinnik K. V. Study of the Method of Classification of News Based on Distributive Semantics	219
Thang Pham Cong, Kopylov A., Thao Tran Thi Thu A Computationally Efficient Algorithm for Image Restoration Using Total Variation	222
Khachumov M. V., Khachumov V. M. On the Safe Achieving the Required Formation Shape by Multiple UAVs (Planar Case)	226
Khrabrov R. N., Fomin M. B. Rendering of the Fluid Flow Through a Thin Walled Pipe with Variable Cross-Section	229
Chistova E. V. Modeling Natural Language Communication with Deep Learning	232
Yadrintsev V. V., Sochenkov I. V. Full Text Patent Classification	235
Yanko J. D. Database Design for Automation of the Voting Process in Application Software	238

Distributed, high-performance computing technologies and big data analytics

Adam Gh., Belyakov D. V., Vala M., Goncharov P. V., Zrellov P. V., Korenkov V. V., Matveyev M. A., Podgainy D. V., Streltsova O. I. Virtual Desktops of the Heterogeneous Cluster HybrILIT	245
Bashashin M. V., Zemlyanaya E. V., Lukyanov K. V. Parallel Implementation of the Microscopic Optical Nucleus-Nucleus Potential Model on the Basis of MPI and OPENMP Technologies	248
Dobrynin V. N., Milovidova A. A., Sokolov I. A. The Assessment of Adequacy of Model to the Object of the Research	252
Dubnov Yu. A. Cross Entropy for Feature Selection in Data Analysis Problems	257
Zuev M. I., Matveev M. A., Podgainy D. V., Streltsova O. I., Torosyan Sh. G. HLTIT-VDI – a New Service of the HybrILIT Ecosystem for Work with Applied Software Packages	260
Mazhitova E. M., Balashov N. A., Baranov A. V., Kutovskiy N. A., Semenov R. N. Integration of the Cloud Infrastructures of the LIT JINR and INP's Astana Branch	262
Nikolsky I. M. Supercomputer Modelling of Wireless Sensor Network in Case of Connection Loss with Base Station	266
Rahmonova A. R. Analysis of the Perfomance of Computer Modeling of the Physical Processes in the System of Long Josephson Junctions in HybrILIT Cluster	269
Chebotov A. I., Avzalov R. L., Chernyshev Y. M., Gertsenberger K. V., Butenko Y. A. Development of BM@N-Webapp Ecosystem for Processing and Visualization of Information about the Experiment on a Fixed Target BM@N	273

Mathematical modeling

Al-Nator M.S., Al-Nator S.V. Optimal Multi-Period Self-Financing Portfolio Strategy with Commission	279
Amirkhanov I.V., Sarker N.R. On a Method of Approximate Solution of the Three-Body Problem	282
Apreutesey A.-M.Y. A Simple Model of Active Queue Management System According to the RED Algorithm	285
Ahmethanova D.M., Zaryadov I.S. Application of Statistical and Probabilistic Methods for Optimization of Parking Systems Operation and for the Recognition of License Plates	288
Baylo N.V., Matyushenko S.I., Pyatkina D.A. Statistical Analysis of Inflation in Contemporary Russia	291
Borovskikh D.A., Dilie GbavleLuk-Donald, Petrov V.A., Savin A.S., Sadovnikov V.K., Khokhlov A.A. Different Applications of SSA Method	294
Bochkarev V.V., Kryanev A.V., Smirnov D.S., Sharafutdinov R.B. Information and Mathematical Model for Supporting Decision-Making on the Decommissioning of Nuclear Power Facilities	297
Budochkina S.A. On Poisson Brackets in Spaces of B_u -potentials	300
Buuruldaï A.E. On the Volatility Function in a Modified Dispersion Model with Constant Elasticity (CEV)	302
Batgerel B., Dimova S.N., Puzynin I.V., Puzynin T.P., Tukhliev Z.K., Hristov I.G., Hristova R.D., Sharipov Z.A. Modeling of Long-Range Thermal Changes of Metal Irradiated by Copper Nanoclusters	305
Velieva T.R., Zavožina A.V., Korolkova A.V. The Determination of the Coefficients of Harmonic Linearization for Deterministic Nonlinear System with Control	308
Velieva T.R., Zavožina A.V., Korolkova A.V. NS-2 as a Verification Means of the Amplitude-Frequency Characteristics of the Control System	311
Gevorkyan M.N., Demidova A.V., Demidova T.S., Sobolev A.A. Simulation of the Spread of Networks Worms with Using Stochastic Differential Equations	314
Gevorkyan M.N., Demidova A.V., Korolkova A.V., Kulyabov D.S., Gostev I.M. Problems in Realization of Stochastic Numerical Runge–Kutta Methods	318
Gogoleva S.Y., Zainetdinova L.G. On an Iterative Method for Solving Ill-Conditioned Systems of Linear Algebraic Equations with Sparse Matrices	321
Grachev D.A., Emelyanova E.U., Popova E.V., Khokhlov A.A. Model of BIA Measurement Hardware	324
Dilie GbavleLuk-Donald, Emelyanova E.U., Popova E.V., Khokhlov A.A. Accelerometer Data Analysis for User Activity Classification	326
Demidova A.V., Druzhinina O.V., Masina O.N., Mijajlovic N., Jacimovic M. Synthesis and Analysis of Multidimensional Mathematical Models of Population Dynamics	329
Denisenko A.P., Mingazhitdinova E.F., Petrov V.A., Savin A.S., Khokhlov A.A. Optimizing Transaction Cost in Ethereum Blockchain	333
Divakov D.V., Sevastianov A.L., Belov M.P. Statement of the Problem of Calculation of the Guided Modes of the Gradient in the Transverse Direction Waveguide	336
Drevitskiy A.S., Divakov D.V., Kotyukov M.M. The Problem of Symbolic-Numeric Computation of the Eigenvalues of the Leaky Modes in a Regular Homogeneous Open Waveguide	339
Egorov A.A. Research of Properties of the Liquid Crystal Waveguide Structures	342
Zotova P.A., Zhukova L.V., Pyatkina D.A. Analysis Vostroyans Popular Formats of Assignments for Online Courses Among Students of Different Educational Institutions	345
Kasimov Y.F., Timerbaev M.I. Optimal Weights of Risk Asset in the Index Portfolio of Constant-Proportion Strategies	348
Kononov M.G., Razumchik R.V. Control of a Discrete-Time Markov Chain with Continuous Bounded State Space which Leads to the Given Stationary Distribution	351
Kryanev A.V., Sliva D.E., Matokhin V.V. Irregularity Measures of Resource Distributions within Economic Systems	355

Kuziv Ia.Iu. The Complex of Programs for Solving Nonlinear Partial Differential Equations of the First Order	358
Kulyabov D.S., Gevorkyan M.N., Korolkova A.V. Numerical Analysis of Geometrized Maxwell Equations	361
Lesik P.A., Romanov A.A., Chistotina D.A. Inverse Problem for the Equation Managementscope	364
Lizhnenko N.M., Morozova D.A. The Problem of Continuation of the Newtonian Potential in a Cylinder of Circular Cross Section	366
Mysina J.A. Rising of Retirement Age and Deciding the Problem of Decreasing the Amount of Employed Population	368
Permyakova J.S. Market Risks of the Commercial Bank: Methods of Estimation and Management	371
Popkov A.Y. Method of Random Vectors Generation with Given Probabilistic Characteristics	374
Pylaeva A.N. Modeling of Price Formation in the Smartphone Market	377
Pyatkina D.A., Matyushenko S.I., Zotova P.A., Kazandaeva A.P. Analysis of Notebook Prices with Econometric Models	381
Rassakhan N.D., Shchetinin E.Yu. Estimation of Heavy Tail Dependence Based on Copulas for the Precipitation Analysis	384
Sevastianov L.A., Vasilyeva D.G. Analysis of Stochastics Dynamic Population Growth	387
Talagaev Y.V. State Estimation of the Nonlinear Network Control Systems Represented by Takagi-Sugeno Fuzzy Models	390
Teveleva E.A., Tretyakov N.P., Terletski A.Ya. Application of the Phase Analysis of Time Series for the Identification of Macroeconomic Cycles Based on the Dynamics of the Exchange Rates	393
Tretyakov N.P., Kochanov A.A. Effective Schrödinger Equation for One-Dimensional Systems with Rapidly Oscillating Boundary Conditions	396
Shebanova K.V., Khamidulina K.I. Stable Calculations of Normal to the Surface with Set Approximately	399
Chupritskiy V.K., Divakov D.V. Simulation of Light Propagation Through Open Planar Waveguide with the Use of “Dirichlet Box” Method	401
Shorokhov S.G., Buuruldai A.E. On Hyperbolic-Sine Local Volatility Model	404
Shumin E.V. Comparison of “Online-Q” and “Experience Replay” for Training a Neural Network in a Multi-Agent Environment	407
On Some Properties of Tail Dependence Coefficient Nonparametric Estimators	410
Shchetinin E.Yu., Berezhkov M.V., Lyubin P.G. Development of an Effective Algorithm for Short-Term Forecasting of Power Consumption Using Ensemble	413
Yu Ying The Method of Finding Algebraic Integral for First-order Differential Equations	416
Babin A.V. Overview of Technologies to Solve the Problems of Monitoring the Performance and Availability of the IT Structure of the Enterprise	419
Konkova E.V. To the Analysis of Phase-type Queuing System in the Special Case	422
Author index	428

Теория телетрафика и её применения
Teletraffic theory and its applications

УДК 519.872

Разработка средства имитационного моделирования ресурсных систем массового обслуживания

К. А. Агеев*, Э. С. Сопин*†

* *Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

† *Институт проблем информатики
Федеральный исследовательский центр «Информатика и управление» РАН
ул. Вавилова, д. 44, к. 2, Москва, Россия, 119333*

Email: kaageyev@gmail.com, sopin_es@rudn.university

Разработано средство имитационного моделирования ресурсных СМО для анализа вероятностных характеристик системы массового обслуживания с ограниченными ресурсами. Результаты анализа применяются для вычисления характеристик М2М трафика.

Ключевые слова: имитационное моделирование, М2М трафик, система массового обслуживания, ограниченные ресурсы.

1. Введение

Системы массового обслуживания (СМО) с ограниченными ресурсами могут быть использованы для анализа показателей современных беспроводных сетей. Основным преимуществом таких СМО является возможность учитывать особенности распределения радио ресурсов в функции распределения требований (ФР) к ресурсам.

Аналитические вычисления вероятностных показателей ресурсных СМО были получены для Пуассоновского поступления заявок в [1, 2]. Но применение таких аналитических формул для вычисления вероятностно-временных характеристик усложняется расчетом сверток ФР требований к ресурсам. В [3] для расчета показателей, в случае дискретного ресурса, предлагается использовать рекуррентный алгоритм расчетов; для случая с непрерывным ресурсом использовался алгоритм дискретизации. Такой подход требует значительных вычислительных мощностей, а алгоритм может быть применен только для Пуассоновского потока поступления заявок. В связи с этим разработка средства имитационного моделирования ресурсных СМО становится важной задачей. Его разработка позволит снизить нагрузку на вычислительные ресурсы при выполнении расчетов вероятностно-временных характеристик.

В работе мы описываем средство имитационного моделирования и его применение для расчетов вероятностно-временных характеристик М2М трафика в беспроводных сетях LTE.

2. Описание средства имитационного моделирования ресурсных СМО

Мы разработали средство имитационного моделирования СМО с ограниченными ресурсами. Алгоритм работы инструмента изображен на рис. 1.

Перед началом имитации происходит инициализация исходных параметров: количество приборов; количество классов заявок; количество типов ресурсов; ограничение по типу ресурсу к классу заявок; времена поступления каждого класса заявок. Далее начинается процесс имитации поступления/обслуживания заявок: выполняется поиск ближайшего события по времени. Если ближайшее событие обслуживания, то выполняется освобождение прибора и ресурса, который занимала заявка. Если поступление, то выполняется проверка свободного прибора. Если свободного прибора нет, то заявка блокируется и пересчитывается статистика. Если свободный прибор есть, то фиксируем класс заявки, время, когда она будет обслужена и объем ресурса, который ей необходим. Затем проверяем достаточно ли для

заявки ресурса: если нет, то заявка заблокирована, пересчитываем статистику и возвращаемся к поиску ближайшего события; если ресурса достаточно, то заявка занимает прибор и выделенный ей ресурс. Затем снова ищем ближайшее событие.

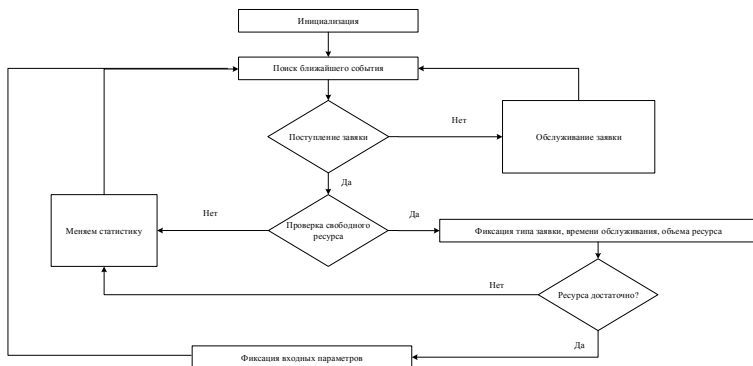


Рис. 1. Схема алгоритма работы инструмента имитационного моделирования

3. Численный эксперимент и анализ результатов

Для проведения анализа показателей качества М2М трафика в беспроводной сети LTE была проведен запуск инструмента имитационного моделирования ресурсной СМО с ФР требований к ресурсам и параметрами из [3, 4].

Имитация проводилась до обслуживания 100 000 заявок. На рис. 2 показано сравнение результатов, полученных при использовании рекуррентного алгоритма и дискретизации [4], и результатов, полученных при имитации.

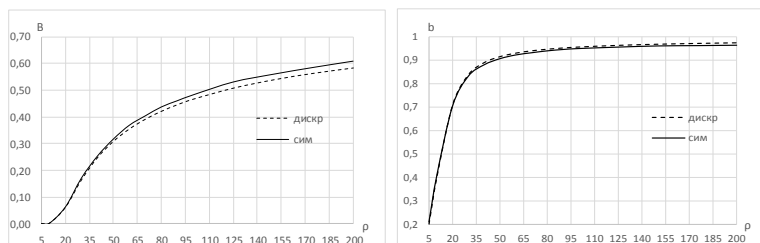


Рис. 2. Сравнение вероятностей блокировки и среднего объема занятого ресурса

Разница в расчетах в процентном соотношении для каждого ρ приведена в табл. 1.

Таблица 1

Погрешность результатов вычисления

ρ	5	20	50	70	90	100	125	150	175	200
B	0,354	2,038	2,976	3,431	3,573	3,686	4,747	4,127	4,139	4,602
b	4,641	0,642	0,968	0,893	0,687	0,689	0,797	0,776	0,981	1,085

4. Заключение

Разработан инструмент имитационного моделирования систем массового обслуживания с ограниченным ресурсом, что позволяет вычислять характеристики M2M трафика в беспроводной сотовой сети учитывая особенности ФР требований к ресурсам, в которых могут учитываться и комбинироваться различные параметры развертывания сети. Расчет этих показателей полезен для последующей конфигурации базовых станций для повышения качества обслуживания абонентов.

Благодарности

Работа частично поддержана грантом РФФИ № 16-07-00766 и № 16-37-60103.

Литература

1. Naumov V. A., Samuilov K. E., Samuilov A. K. On the total amount of resources occupied by serviced customers // Autom Remote Control. — 2016. — Vol. 77, issue 8. — Pp. 1419–1427.
2. Tikhonenko O., Generalized Erlang problem for service systems with finite total capacity // Problems of Information Transmission. — 2005. — 41 (3). — Pp. 243–253.
3. Sopin E., Markova E., Vikhrova O., Ageev K., Gudkova I., Samouylov K., Performance analysis of M2M traffic in LTE network using queuing systems with random resource requirements // Automatic Control and Computer Sciences (in print).
4. Samouylov K., Sopin E., Vikhrova O. On design of efficient algorithm for blocking probability calculation in queuing system with random requirements // 15th International Scientific Conference, ITMM 2016. — September 12–16, 2016.

UDC 519.872

Development of the Simulation Tool for Resource Queuing Systems

K. A. Ageev*, E. S. Sopin*[†]

* Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation

[†] Institute of Informatics Problems, FRC CSC RAS
44-2 Vavilova Str., Moscow 119333, Russian Federation

Email: kaageyev@gmail.com, sopin_es@rudn.university

Simulation tool for resource queuing systems for analysis of M2M performance characteristics of queuing system with limited resources was built. The results can be applied for computing performance metrics of M2M traffic.

Key words and phrases: M2M traffic, simulation model, queuing system; limited resources.

УДК 519.872

Рекуррентный алгоритм для расчета показателей эффективности СМО с приборами, находящимися под влиянием случайной среды

К. И. Аду, Е. В. Маркова

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: 1032135491@pfur.ru, markova_ev@pfur.ru

Построена модель одной соты беспроводной сети связи четвертого поколения, представленная в виде системы массового обслуживания, приборы которой находятся в случайной среде, способной оказывать влияние на их состояние. Получен рекуррентный алгоритм для расчета вероятности блокировки.

Ключевые слова: СМО, случайная среда, прерывание обслуживания, рекуррентный алгоритм, вероятность блокировки.

1. Введение

Рассмотрим некоторую случайную среду, способную оказывать влияние на показатели качества обслуживания пользователей в рамках одной соты беспроводной сети передачи данных, радиус которой равен R . Предположим, что все пользователи соты находятся на одинаковом расстоянии от БС, определим это расстояние как d . В систему поступают запросы пользователей, генерирующие потоковый трафик [1], характеризующийся фиксированной скоростью передачи данных r_0 .

Под влиянием случайной среды мощность передачи данных может меняться, что, в свою очередь, отражается на скорости передачи данных и на максимальном количестве обслуживаемых пользователей. Далее перейдем к подробному описанию частного случая модели, в котором мощность передачи данных может быть изменена только дважды. Для расчета вероятностных характеристик этого случая получен рекуррентный алгоритм.

2. Построение математической модели

Входящий поток запросов имеет пуассоновское распределение с параметром λ . Среднее время обслуживания одного запроса равно μ^{-1} . Случайная среда может находиться в двух состояниях $s = 0, 1$, где $s = 1$ означает передачу данных с максимальной мощностью p_1 , $s = 0$ — передачу данных с минимальной мощностью p_0 .

Под действием внешней среды происходит снижение мощности передачи данных [2] с максимального значения p_1 до минимального значения p_0 при этом среда с интенсивностью α переходит из состояния $s = 1$ в состояние $s = 0$, с интенсивностью β происходит обратный переход и восстановление мощности до первоначального уровня p_1 .

При переходе среды в состояние $s = 0$ происходит снижение числа обслуживаемых запросов пользователей с N_1 до N_0 . Максимальное число N_s , $s = 0, 1$ (1) обслуживаемых в системе пользователей может быть рассчитано как отношение достижимой скорости передачи данных r_s , $s = 0, 1$, определяемой согласно формуле Шеннона [3] и в соответствии с текущим состоянием среды, к гарантированной скорости r_0 .

$$N_s = \left\lfloor \frac{\omega \ln \left(1 + \frac{G p_s}{d^\kappa N_0} \right)}{r_0} \right\rfloor, \quad s = 0, 1. \quad (1)$$

где ω — ширина полосы частот канала, G — коэффициент затухания сигнала, κ — степень затухания сигнала, N_0 — мощность шума.

Так как $N_1 > N_0$, под воздействием случайной среды происходит прерывание обслуживания $N_1 - N_0$ запросов. Состояние модели описывает двумерный вектор (n, s) над пространством состояний $X = \{(n, s): 0 \leq n \leq N_s, s = 0, 1\}$, где n — число обслуживаемых запросов пользователей, s — состояние случайной среды.

Элементы матрицы интенсивностей переходов для данной системы определяются следующим образом (2):

$$a((n, s)(n', s')) = \begin{cases} \alpha, & \text{если } n' = n, s' = s - 1, n = \overline{0, N_{s-1}}, s = 1, \\ & \text{или } n' = N_{s-1}, s' = s - 1, n \geq N_{s-1}, s = 1; \\ \beta, & \text{если } n' = n, s' = s + 1, n = \overline{0, N_s}, s = 0; \\ \lambda, & \text{если } n' = n + 1, s' = s, n < N_s; \\ n\mu, & \text{если } n' = n - 1, s' = s, n > 0; \\ *, & \text{если } n' = n, s' = s; \\ 0, & \text{в противном случае,} \end{cases} \quad (2)$$

где $*$ = $-(\alpha \cdot I(s = 1) + \beta \cdot I(s = 0, n \leq N_s) + \lambda \cdot I(n < N_s) + n\mu \cdot I(n > 0))$, а I — функция индикатор.

3. Рекуррентный алгоритм

Для расчета стационарного распределения вероятностей $p(n, s)$, $(n, s) \in X$, был получен рекуррентный алгоритм. Введем ненормированные вероятности $q(n, s)$, которые вычисляются по формулам (3) и (4)

$$q(0, 0) = 1, \quad q(0, 1) = x; \quad (3)$$

$$q(n, s) = \gamma_{ns} + \delta_{ns}x, \quad (n, s) \in X : n \geq 0. \quad (4)$$

В формуле (4) коэффициенты γ_{ns} и δ_{ns} вычисляются по рекуррентным соотношениям (5)

$$\left\{ \begin{array}{l} \gamma_{00} = 1, \delta_{00} = 0; \quad \gamma_{01} = 0, \delta_{01} = 1; \\ \gamma_{10} = \frac{\lambda + \beta}{\mu}, \quad \delta_{10} = -\frac{\alpha}{\mu}; \quad \gamma_{11} = -\frac{\beta}{\mu}, \quad \delta_{11} = -\frac{\lambda + \alpha}{\mu}; \\ \gamma_{n1} = \frac{1}{n} [(\rho + (n - 1) - \delta_{10})\gamma_{n-1,1} - \rho\gamma_{n-2,1} - \gamma_{11}\gamma_{n-1,0}], \quad n = \overline{2, N_0}, \\ \delta_{n1} = \frac{1}{n} [(\rho + (n - 1) - \delta_{10})\delta_{n-1,1} - \rho\delta_{n-2,1} - \gamma_{11}\delta_{n-1,0}], \quad n = \overline{2, N_0}; \\ \gamma_{n0} = \frac{1}{n} [(\gamma_{10} + (n - 1))\gamma_{n-1,0} + \delta_{10}\gamma_{n-1,1} - \rho\gamma_{n-2,0}], \quad n = \overline{2, N_0 - 1}, \\ \delta_{n0} = \frac{1}{n} [(\gamma_{10} + (n - 1))\delta_{n-1,0} + \delta_{10}\delta_{n-1,1} - \rho\delta_{n-2,0}], \quad n = \overline{2, N_0 - 1}; \\ \gamma_{n1} = \frac{1}{n} [(\rho + (n - 1) - \delta_{10})\gamma_{n-1,1} - \rho\gamma_{n-2,1}], \quad n = \overline{N_0 + 1, N_1 - 1}, \\ \delta_{n1} = \frac{1}{n} [(\rho + (n - 1) - \delta_{10})\delta_{n-1,1} - \rho\delta_{n-2,1}], \quad n = \overline{N_0 + 1, N_1 - 1}. \end{array} \right. \quad (5)$$

Используя систему уравнений равновесия, определим x (6):

$$x = \frac{-(\alpha + N_1\mu)\gamma_{N_1,1} + \lambda\gamma_{N_1-1,1}}{(\alpha + N_1\mu)\delta_{N_1,1} - \lambda\delta_{N_1-1,1}}. \quad (6)$$

Стационарное распределение вероятностей состояний модели определяется по формуле $p(n, s) = q(n, s) / \sum_{(i, j) \in X} q(i, j)$, $(n, s) \in X$.

Основной вероятностной характеристикой модели является вероятность блокировки, которая рассчитывается по формуле $B = p(N_0, 0) + p(N_1, 1)$.

4. Заключение

В ходе данной работы была рассмотрена модель одной соты беспроводной сети связи четвертого поколения, представленная в виде СМО, находящейся в случайной среде. Исследован частный случай модели. Предложен рекуррентный алгоритм для расчета вероятности блокировки запросов на обслуживание пользователей.

Целью дальнейших исследований является решение задачи оптимизации количества приборов в случайной среде по критерию максимизации числа обслуживаемых запросов и минимизации числа прерываемых пользователей.

Благодарности

Публикация подготовлена при финансовой поддержке РФФИ в рамках научного проекта № 18-37-00231.

Литература

1. Башарин Г. П. Лекции по математической теории телеграфика: Учеб. пособие. Изд. 3-е, испр. и доп. — М.: РУДН, 2009. — 342 с.
2. Ahmadian A., Galinina O., Gudkova I., Andreev S., Shorgin S., and Samouylov K. On capturing spatial diversity of joint M2M/H2H dynamic uplink transmissions in 3GPP LTE cellular system // *Lecture Notes in Computer Science*. — 2014. — Vol. 9247. — Pp. 407–421.
3. Markova E., Gudkova I., Ometov A., Dzantiev I., Andreev S., Koucheryavy Ye., Samouylov K. Flexible Spectrum Management in a Smart City within Licensed Shared Access Framework // *IEEE Access*. — 2017. — Vol. 5. — Pp. 22252–22261.

UDC 519.872

Recursive Algorithm for Calculating Indicators of Queuing System Model with Devices under Influence of a Random Environment

K. Y. Adou, E. V. Markova

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: 1032135491@pfur.ru, markova_ev@pfur.ru

In this report, the model of one cell of a wireless network in the fourth generation is constructed. The model is presented as a queuing system, where devices are in a random environment, which influences their state. A recursive algorithm is obtained to calculate the blocking probability.

Key words and phrases: queuing system, random environment, service interruption, recursive algorithm, blocking probability.

УДК 519.872.5

Анализ модели процесса обслуживания клиентов страховой компании с ограниченным сроком действия договоров страхования

Н. В. Багаева

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: nvbagaeva@mail.ru

Разнообразие областей, которые позволяют исследовать современный математический аппарат, с каждым годом стремительно увеличивается. Методы теории массового обслуживания, существующие уже на протяжении века, ориентированы на анализ систем (СМО) и сетей массового обслуживания (СеМО). Практическому применению способствуют простота и естественность, с которыми они отображают структуры моделируемых систем и процессы обработки в системах объектов различных типов. В статье рассматривается модель процесса обслуживания клиентов страховой компании с ограниченным сроком действия договоров страхования, которой соответствует открытая сеть (ОСеМО), имеющая пуассоновский поток заявок и дисциплину обслуживания First Come First Served (FCFS).

Ключевые слова: страховая компания, договор страхования, иск, сеть массового обслуживания, доход.

1. Введение

В настоящий момент большой интерес вызывают модели актуарной математики, т.е. математики, которая связана со страховым делом. В числе этих вопросов находится проблема математической модели процессов функционирования страховой компании. К сожалению, в работе [1, С. 137–146] рассмотрена модель страхового договора, имеющая ограничения на количество клиентов, не дающая возможность рассмотреть ситуацию максимально приближенную к реальности. Ей соответствует замкнутая сеть. В данной работе построена модель, предусматривающая произвольное число клиентов k , $k = \overline{0, \infty}$. А также применены способы исследования открытой сети, полученные в [3, С. 276–278]. При описании маршрутов заявок используется теория марковских блужданий [2, С. 500–511]. В исследуемой модели рассматривается экспоненциальное обслуживание.

2. Постановка задачи

Задача анализа модели страхового договора с ограниченным сроком страхования и неограниченным числом клиентов является новой. Для составления модели подробно были изучены виды договоров и типы страхования [4, с. 124–185]. Рассматривается процесс функционирования СК, заключающей договоры страхования $(n - 1)$ -го типов:

S_1 — договор краткосрочного страхования;

⋮

S_{n-1} — договор страхования на дожитие.

Возможное число клиентов компании равно k , $k = \overline{0, \infty}$. Заключением договором занимаются m_{n+2} сотрудников компании — страховщики. При предъявлении иск проходит две стадии обработки: стадию оценки и стадию выплаты. Оценкой исков типа i занимается m_i сотрудников компании — оценщики, $i = \overline{1, n-1}$. Выплатой по предъявленным и уже оцененным искам, а также приемом платежей по заключенным договорам страхования занимаются m_{n+1} кассиров (рис. 1).

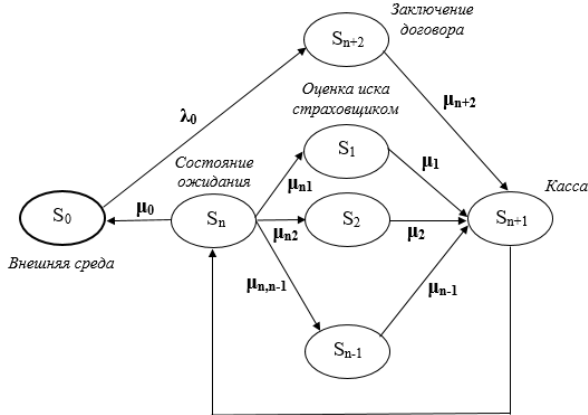


Рис. 1. Структура сети

Каждый из клиентов компании может находиться в одном из следующих состояний: C_n — в состоянии ожидания, предъявлять иск не требуется; C_i — на стадии оценки иска i -го типа; $i = \overline{1, n-1}$; C_{n+2} — на этапе осуществления кассовых операций; C_{n+1} — на этапе заключения договора; C_0 — пребывания потенциально возможного клиента во «внешней среде».

Состояние СК в момент времени t опишем с помощью вектора (1):

$$k(t) = (k, t) = (k_1(t), k_2(t), \dots, k_{n-1}(t), k_n(t), k_{n+1}(t), k_{n+2}(t)), \quad (1)$$

где $k_i(t)$ — количество исков типа i , находящихся в состоянии C_i , $i = \overline{1, n-1}$; $k_j(t)$ — общее число заявок, находящихся в состоянии C_j , $j = \overline{n, n+2}$.

Времена обработки исков оценщиками, обслуживания клиентов кассирами и заключения договоров страховщиками распределены по показательному закону с параметром μ_i , μ_{n+1} , μ_{n+2} соответственно, $i = \overline{1, n-1}$. Переход клиента из состояния C_0 в состояние C_{n+2} происходит в случайные моменты времени, независимо от того, на каком этапе обслуживания находятся другие клиенты, и независимо от времени.

В ходе работы была выведена система уравнений равновесия для данной модели:

$$p(\vec{k}) \left[\lambda_0 + \sum_{i=1}^{n-1} \mu_i(k_i) + \sum_{i=1}^{n-1} \mu_{ni}(k_n) + \mu_0 k_n + \mu_{n+1} k_{n+1} + \mu_{n+2} k_{n+2} \right] = \\ = p(\vec{k} - \vec{e}_{n+2}) \cdot \lambda_0 \cdot u(k_{n+2} - 1) + p(\vec{k} - \vec{e}_{n+1} + \vec{e}_{n+2}) \cdot \mu_{n+2}(k_{n+2} + 1) \cdot u(k_{n+1} - 1) +$$

$$\begin{aligned}
&= p(\vec{k} + \vec{e}_{n+1} - \vec{e}_n) \cdot \mu_{n+1}(k_{n+1} + 1) \cdot u(k_n - 1) + p(\vec{k} + \vec{e}_n) \cdot \mu_0(k_n + 1) + \\
&+ \sum_{i=1}^{n-1} p(\vec{k} + \vec{e}_n - \vec{e}_i) \cdot \mu_{ni}(k_n + 1) \cdot u(k_i - 1) + \sum_{i=1}^{n-1} p(\vec{k} + \vec{e}_i - \vec{e}_{n-1}) \cdot \mu_i(k_i + 1) \cdot u(k_{n-1} - 1).
\end{aligned}$$

3. Заключение

Рассмотрена многолинейная экспоненциальная сеть. Её узлами являются СМО с пуассоновским входящим потоком, неограниченным накопителем, с пуассоновским распределением времени обслуживания заявки и дисциплиной обслуживания FCFS. Выписаны основные характеристики модели. А также для нее получена и решена СУР.

Литература

1. Русилко Т., Матылицкий М. Сетевые вероятностные модели обработки заявок в страховых компаниях, 2012.
2. Бочаров П. П., Печенкин А. В. Теория массового обслуживания. — М.: РУДН, 1995.
3. Митрофанов Ю. И. Теория массового обслуживания, 1995.
4. Иванова Л. В., Звездина Н. В. Актуарные расчеты в страховании жизни и пенсионном страховании, 2012.

UDC 519.872.5

The Analysis of Models of Insurance Contracts with a Limited Term of Service

N. V. Bagaeva

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: nvbagaeva@mail.ru

The diversity of fields that allows you to explore advanced mathematical apparatus, every year is rapidly increasing. Methods of the queueing theory, existing for over a century, focused on the analysis of systems and queueing networks. The practical application contributes to the simplicity and naturalness with which they display the structure of the simulated systems and processes in systems of objects of several types. In the article, I consider the model of insurance contract with a limited insurance period, which corresponds to an open network with Poisson flow of requests, a queueing discipline First Come First Served (FCFS).

Key words and phrases: insurance company, insurance policy, lawsuit, queueing network, income.

УДК 519.872, 519.217

Математическая модель системы с рекуррентным входящим потоком, повторным обслуживанием и обобщенным обновлением

Е. В. Богданова*, И. С. Зарядов*†, Т. А. Милованова*

* Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198

† Институт проблем информатики
Федеральный исследовательский центр «Информатика и управление» РАН
ул. Вавилова, д. 44, к. 2, Москва, Россия, 119333

Email: official_kb@mail.ru, zaryadov_is@rudn.university, milovanova_ta@rudn.university

Строится математическая модель системы с рекуррентным входящим потоком, с повторным обслуживанием и механизмом обобщенного обновления (сбросом принятых заявок из накопителя).

Ключевые слова: система массового обслуживания, обобщенное обновление, повторное обслуживание, вероятностные характеристики.

1. Введение

Построение и исследование математических моделей систем с возможной потерей по тем или иным причинам (ненадежный прибор — [1], поступление заявок особого типа — [2,3], сброс принятых в систему заявок обслужившейся на приборе заявкой — [4–11]) является актуальной задачей.

В работе [4] впервые был введен механизм обновления, когда заканчивающая обслуживание на приборе заявка может с вероятностью q сбросить все заявки из накопителя и покинуть систему, либо с дополнительной вероятностью $p = 1 - q$ просто покидает систему. В работе [5] П. П. Бочаровым было введено обновление с повторным обслуживанием, когда обслужившаяся заявка опустошает накопитель и остается в системе для повторного обслуживания.

В дальнейшем модели с обновлением были развиты А. В. Печинкиным, предложившим рассматривать случай, когда заявка в момент ухода с прибора с вероятностью $q(i)$ сбрасывает из накопителя ровно i заявок ($i \geq 1$) [6,7] — обобщенное обновление.

Если ранее, как правило, рассматривались системы типа $M|M|n|r$ или $GI|M|n|r$ [4–7], то в [8,9] уже исследовалась система вида $M|G|1|r$.

Параллельно изучались системы с повторным обслуживанием и обновлением. Первая работа — [5], в [10,11] достаточно полно была рассмотрена система вида $GI|M|1|\infty$ с обновлением (накопитель опустошается полностью) и повторным обслуживанием, были выведены основные вероятностно-временные характеристики системы.

Первая попытка перенести механизм обобщенного обновления на случай повторного обслуживания была сделана в [12] для системы $M|M|1|\infty$, были получены в явном виде основные характеристики функционирования этой системы. В [13] были сформулированы возможные подходы к исследованию системы $GI|M|1|\infty$.

Задача данной работы — продолжить, начатое в [13] (с учетом найденных ошибок) и рассмотреть обобщенное обновление с повторным обслуживанием для системы типа $GI|M|1|\infty$.

2. Описание модели

Рассматривается система $GI|M|1|\infty$ с рекуррентным входящим потоком, экспоненциальным обслуживанием заявок на приборе, накопителем неограниченной емкости и механизмом обобщенного обновления с повторным обслуживанием.

Обобщенное обновление с повторным обслуживанием определяется следующим образом. С вероятностью $q(i)$ ($i \geq 0$) находящаяся на приборе заявка в момент окончания обслуживания сбрасывает из накопителя ровно i заявок, если в накопителе больше i заявок (если меньше — то накопитель опустошается полностью), и остается в системе. Возможно два варианта повторного обслуживания: либо обслуживавшаяся заявка занимает первое свободное место в накопителе (рассматриваемый вариант), либо остается на приборе для повторного обслуживания. С вероятностью p обслуживавшаяся заявка просто покидает систему. Стоит отметить,

что $\sum_{i=0}^{\infty} q(i) + p = 1$.

Для исследования строится вложенная по моментам поступления цепь Маркова (ЦМ) $\nu_n = \nu(\tau_n - 0)$ (τ_n — момент поступления n -й заявки в систему) со счетным множеством состояний $\mathcal{X} = \{0, 1, 2, \dots\}$ и матрицей переходных вероятностей $P = (p_{ij})_{i,j \geq 0}$. Для нахождения стационарного по вложенной ЦМ распределения p_k^- , $k \geq 0$ вводятся вспомогательные вероятности: $\pi(k, n, m)$ — вероятность того, что за время между последовательными моментами поступления заявок на приборе обслужится ровно k ($k \geq 1$) заявок, из накопителя будет сброшено ровно m ($m \geq 0$) заявок, и из k обслужившихся заявок ровно n ($0 \leq n \leq k - 1$) покинут систему, при условии, что в момент времени $\tau_n - 0$ в системе было не менее $n + m$ заявок; $\pi^*(k, n, m)$ — вероятность того, что k ($k \geq 1$) обслужившихся на приборе заявок полностью опустошат накопитель, если в нем изначально было m ($m \geq 0$) заявок, и из k обслужившихся заявок ровно n ($0 \leq n \leq k - 1$) покинут систему. С помощью этих вспомогательных вероятностей однозначно определяются элементы матрицы P переходных вероятностей вложенной цепи Маркова.

Данное стационарное распределение имеет геометрический вид, аналогично [6, 7, 10, 11].

3. Заключение

Работа представляет начальную часть исследования системы $GI|M|1|\infty$ с обобщенным обновлением и повторным обслуживанием. Получены основные вероятностные характеристики.

Благодарности

Публикация подготовлена при поддержке Программы РУДН «5-100» и при финансовой поддержке РФФИ в рамках научных проектов № 18-07-00692, 16-07-00766.

Литература

1. Dudin A., Klimenok V., Vishnevsky V. Analysis of Unreliable Single Server Queueing System with Hot Back-Up Server // Communications in Computer and Information Science. — 2015. — Vol. 499. — Pp. 149–161.
2. Gelenbe E. Product-form queueing networks with negative and positive customers // Journal of Applied Probability. — 1991. — Vol. 28 (3). — Pp. 656–663.
3. Бочаров П. П., Д'Апиче Ч., Манзо Р., Печинкин А. В. Анализ многолинейной марковской системы массового обслуживания с неограниченным накопителем

- и отрицательными заявками // Автоматика и телемеханика. — 2007. — №. 1. — С. 93–104.
4. Kreinin A. Queueing Systems with Renovation // Journal of Applied Math. Stochast. Analysis. — 1997. — Vol. 10, no. 4. — Pp. 431–443.
 5. Бочаров П. П., Зарядов И. С. Стационарное распределение вероятностей в системах массового обслуживания с обновлением // Вестник РУДН. Серия: Математика. Информатика. Физика. — 2007. — № 1–2. — С. 14–23.
 6. Зарядов И. С., Печинкин А. В. Стационарные временные характеристики системы $G|M|n|\infty$ с некоторыми вариантами дисциплины обобщённого обновления // Автоматика и телемеханика. — 2009. — № 12. — С. 161–174.
 7. Zaryadov I. S., Razuimchik R. V., Milovanova T. A. Stationary waiting time distribution in $G/M/n/r$ with random renovation policy // Communications in Computer and Information Science. — 2016. — Vol. 678. — Pp. 349–360.
 8. Bogdanova E. V., Milovanova T. A., Zaryadov I. S. The Analysis of Queueing System with General Service Distribution and Renovation // Вестник РУДН. Серия: Математика. Информатика. Физика. — 2017. — Т. 25, № 1. — С. 3–8.
 9. Zaryadov I. S., Bogdanova E. V., Milovanova T. A. Probability-time characteristics of $M|G|1|1$ queueing system with renovation // CEUR Workshop Proceedings. 2017. — Vol. 1995. — Pp. 125–131.
 10. Зарядов И. С., Королькова А. В., Милованова Т. А., Щербанская Т. А. Математическая модель расчета и анализа характеристик систем с обобщенным обновлением и повторным обслуживанием // Т-Сопм: Телекоммуникации и транспорт. — 2014. — № 6. — С. 16–20.
 11. Zaryadov I. S., Scherbanskaya A. A. Time Characteristics of Queueing System with Renovation and Reserve // Вестник РУДН. Серия: Математика. Информатика. Физика. — 2014. — № 2. — С. 61–66.
 12. Мацкевич И. А., Вероятностно-временные характеристики системы массового обслуживания с обобщенным обновлением и повторным обслуживанием // Материалы Всероссийской конференции с международным участием «Информационно-телекоммуникационные технологии и математическое моделирование высоко-технологичных систем». — М.: РУДН, 2016. — С. 37–39.
 13. Зарядов И. С., Мацкевич И. А., Щербанская А. А. Система массового обслуживания с обобщенным обновлением и повторным обслуживанием — вероятностно-временные характеристики // Материалы Девятнадцатой международной научной конференции «Распределенные компьютерные и телекоммуникационные сети: управление, вычисление, связь (DCCN-2016)». — 2016. — Т. 3. — С. 458–462.

UDC 519.872, 519.217

The Mathematical Model of the Retrial Queueing System with General Renovation and Recurrent Input Flow

E. V. Bogdanova*, I. S. Zaryadov*[†], T. A. Milovanova*

* Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation

[†] Institute of Informatics Problems, FRC CSC RAS
44-2 Vavilova str., Moscow 119333, Russian Federation

Email: official_kb@mail.ru, zaryadov_is@rudn.university, milovanova_ta@rudn.university

The mathematical model of the system with recurrent input flow, repeated service and general renovation (tasks from the buffer may be dropped out) is considered.

Key words and phrases: queueing system, general renovation, recurrent input flow, repeated service, retrial queue, probability characteristics.

УДК 519.872

Модель соты беспроводной M2M сети с равномерно распределенными на плоскости источниками заявок в виде СМО

А. А. Гольская, Е. В. Маркова, И. Л. Дзантиев

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: feduro.nastya@yandex.ru, markova_ev@pfur.ru, dzonthammet@mail.ru

Важным параметром при определении показателей эффективности беспроводных сетей межмашинного взаимодействия (Machine-to-Machine, M2M) – скорости передачи данных, вероятности блокировки – стала удаленность устройства от базовой станции (БС). При описании такой сети в виде системы массового обслуживания (СМО) с потоковым или эластичным трафиком необходимо учесть расстояние между устройством и БС. В работе построена модель соты беспроводной сети со стационарными M2M устройствами, описываемыми точками, случайно возникающими на плоскости.

Ключевые слова: LTE, беспроводная сеть, межмашинное взаимодействие, M2M, индикатор качества канала, CQI, вероятность блокировки, управление радиоресурсами.

1. Введение

Согласно прогнозу компании Cisco Systems [1], к 2021 году количество M2M соединений вырастет до 3,3 миллиардов, что может привести к нехватке частотного диапазона, для эффективного использования которого могут быть применены различные политики управления частотно-временными ресурсами. Единица частотно-временного ресурса – ресурсный блок, параметрами которого являются время, частота и мощность, согласно рекомендации 3GPP характеризуется также индикатором качества канала CQI (Channel Quality Indicator), который может принимать значения $c = \overline{1, L}, L = 15$. Чем выше значение – тем выше скорость, которую может выделить БС.

В работе построена модель соты беспроводной сети со стационарными M2M устройствами, находящимися в пассивном или активном состоянии, описываемыми точками, случайно возникающими на плоскости. Устройства генерируют потоковый трафик, передаваемый с гарантированной скоростью r_0 , при этом максимальная достижимая скорость зависит от удаленности устройства от БС, мощности передачи сигнала, уровня шума. Рассмотрена одна из возможных дисциплин разделения радиоресурсов – «full power», при которой передача данных осуществляется с максимальной мощностью [2].

2. Построение математической модели

Рассмотрим соту сети беспроводной передачи данных радиуса R с равномерно распределенными стационарными M2M устройствами. Устройства с интенсивностью λ переходят в активное состояние и передают данные на максимальной мощности p_{\max} согласно используемой политике «full power».

Пусть случайная величина (СВ) $\xi_d(\eta) = RL^{-1}\eta$, где $\eta = L + 1 - c$, определяет максимальное расстояние от устройства до БС. Обозначим минимальное расстояние $\xi_d(1)$ как d_0 . Согласно формуле Шеннона достижимая M2M устройством скорость передачи данных $r(\xi_d(\eta), p_{\max})$ зависит от ширины полосы частот ω , мощности передачи сигнала p_{\max} , удаленности $\xi_d(\eta)$ от БС и определяется по формуле (1):

$$r(\xi_d(\eta), p_{\max}) = \omega \ln \left(\frac{1 + Gp_{\max}}{\xi_d^{\kappa}(\eta) N_0} \right) = \omega \ln \left(\frac{1 + Gp_{\max}}{\left(\frac{R}{L}\eta\right)^{\kappa} N_0} \right), \quad (1)$$

где G — коэффициент затухания сигнала, N_0 — уровень шума, а κ — степень затухания.

Так как M2M устройства распределены в сети равномерно, то функция распределения (ФР) СВ $\xi_d(\eta)$ имеет вид (2):

$$F_{\xi_d(\eta)}(d) = P\{\xi_d(\eta) \leq d\} = \begin{cases} 0, & \text{при } d < 0; \\ \frac{d^2}{R^2}, & \text{при } d \in [0, R]; \\ 1, & \text{при } d > R, \end{cases} \quad (2)$$

а соответствующая плотность распределения $f_{\xi_d(\eta)}(d) = \frac{2d}{R^2}$, $0 \leq d \leq R$.

Обозначим $P(k-1)$ условную вероятность того, что k -е M2M устройство с соответствующим значением индикатора CQI, определяемым при помощи СВ η_k , которая принимает значения l_k , $l_k \in \{1, \dots, L\}$, будет обслужено, при условии, что в системе уже обслуживается $k-1$ M2M устройство, каждое из которых имеет соответствующее значение индикатора CQI l_1, \dots, l_{k-1} . Тогда диаграмма интенсивностей переходов будет иметь вид, представленный на рис. 1.

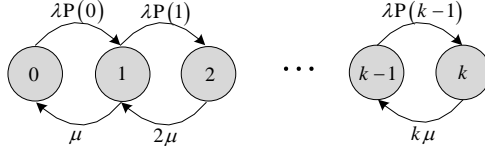


Рис. 1. Диаграмма интенсивностей переходов

Если число обслуживаемых устройств равно нулю, вероятность того, что новое устройство будет принято на обслуживание, вычисляется по формуле (3):

$$P(0) = F_{\xi_d(\eta)}\left(\min\left\{R, \left(Gp_{\max}/(e^{r_0/\omega} - 1)N_0\right)^{1/\kappa}\right\}\right). \quad (3)$$

Условная вероятность $P(k)$ того, что $k+1$ -е M2M устройство будет обслужено в системе с k активными устройствами при применении политики «full power» определяется по формуле (4):

$$P(k) = \frac{P\left\{\sum_{i=1}^{k+1} \frac{1}{r(l_i, p_{\max})} \leq \frac{1}{r_0}, \sum_{i=1}^k \frac{1}{r(l_i, p_{\max})} \leq \frac{1}{r_0}\right\}}{P\left\{\sum_{i=1}^k \frac{1}{r(l_i, p_{\max})} \leq \frac{1}{r_0}\right\}}, \quad (4)$$

где согласно центральной предельной теореме $P\left\{\sum_{i=1}^k \frac{1}{r(l_i, p_{\max})} \leq \frac{1}{r_0}\right\} = \Phi\left(\frac{1 - \theta k r_0}{r_0 \sigma \sqrt{k}}\right)$, θ и σ — математическое ожидание и среднеквадратическое отклонение независимых и одинаково распределённых СВ $\frac{1}{r(l_i, p_{\max})}$, $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp^{-\frac{t^2}{2}} dt$ — стандартное нормальное распределение.

Для сокращения дальнейших расчетов введем m_k и τ_k^2 :

$$m_k = k r_0 E\left[\frac{1}{r(l_i, p_{\max})}\right], \tau_k^2 = k r_0 \left(E\left[\left(\frac{1}{r(l_i, p_{\max})}\right)^2\right] - \left(E\left[\frac{1}{r(l_i, p_{\max})}\right]\right)^2\right),$$

где $E[1/r(l_i, p_{\max})] = (1/r(d_0, p_{\max})) F_{\xi_d(\eta)}(d_0) + \int_{d_0}^R \frac{1}{\omega \ln(1 + G p_{\max}/x^k N_0)} f_{\xi_d(\eta)}(x) dx$
 $E\left[\left(\frac{1}{r(l_i, p_{\max})}\right)^2\right] = (1/r(d_0, p_{\max}))^2 F_{\xi_d(\eta)}(d_0) + \int_{d_0}^R \frac{1}{\omega^2 \ln^2(1 + G p_{\max}/x^k N_0)} f_{\xi_d(\eta)}(x) dx.$
 Тогда условная вероятность $P(k)$ определяется по формуле (5):

$$P(k) = \Phi\left(\frac{1 - m_{k+1}}{\tau_k + 1}\right) / \Phi\left(\frac{1 - m_k}{\tau_k}\right). \quad (5)$$

3. Заключение

В работе представлена модель соты беспроводной сети с M2M устройствами, равномерно распределенными в соте и сгруппированными по значению индикатора CQI. Устройства генерируют потоковый трафик, передаваемый с гарантированной скоростью. Рассмотрена политика управления радиоресурсами «full power».

Благодарности

Публикация подготовлена при финансовой поддержке РФФИ в рамках научного проекта № 18-37-00231.

Литература

1. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021 White Paper.
2. Ahmadian A., Galinina O., Gudkova I., Andreev S., Shorgin S., Samouylov K. On capturing spatial diversity of joint M2M/H2H dynamic uplink transmissions in 3GPP LTE cellular system, Lecture Notes in Computer Science. — 2014. — Vol. 9247. — Pp. 407–421.

UDC 519.872

Scheme Model of Wireless M2M Network with Uniformly Distributed Request Sources in the Plane

A. A. Golskaia, E. V. Markova, I. L. Dzantiev

*Department of Applied Probability and Informatics
 Peoples' Friendship University of Russia (RUDN University)
 6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: feduro.nastya@yandex.ru, markova_ev@pfur.ru, dzonthammet@mail.ru

An important parameter the performance metrics determination of wireless Machine-to-Machine networks (M2M) is the farness of the device from the base station (BS). In case of describing such a network as a queuing system with streaming or elastic traffic, it is necessary to take into account the distance between the device and the BS. The model of a wireless network cell with stationary M2M devices, which is described by points randomly occurring on a plane, is constructed.

Key words and phrases: LTE, wireless network, Machine-to-Machine communication, M2M, channel quality indicator, CQI, blocking probability, radio resource management.

УДК 004.032.26

Об исследовании параметров метеорологических моделей на основе паттернов

А. К. Горшенин

Федеральный исследовательский центр «Информатика и управление» РАН
ул. Вавилова, д. 44, к. 2, Москва, Россия, 119333

Email: agorshenin@frccsc.ru

В работе предложено использование цепочек событий (паттернов), выделенных из данных, для анализа процессов выпадения осадков. Исследованы вопросы марковости данных, построения вероятностных и нейростевых прогнозов для базовых статистических наблюдений без привлечения каких-либо дополнительных сведений о метеорологических условиях. Наилучшая достигнутая точность нейросетей, обучаемых на паттернах, составила 97% для однодневных и 89% для двухдневных прогнозов.

Ключевые слова: осадки, паттерны, прогноз, нейронные сети, глубокое обучение, вероятностное прогнозирование.

1. Введение

Осадки являются важным параметром для метеорологических моделей (см., например, статью [1]), поэтому востребованы разработка адекватного математического аппарата (в том числе вероятностно-статистического) и создание программных инструментов для обработки значительного объема накопленных наблюдений с использованием современных методов. При этом для решения задач прогнозирования могут быть использованы как вероятностные подходы (см., например, работу [2]), так и весьма эффективные в широком спектре прикладных областей нейронные сети (см., например, статью [3]). Более того, в настоящее время достаточно популярной является тематика исследования различных процессов выпадения осадков в контексте задач глобального потепления и изменения климата (см., например, работу [4]). В данной работе коротко рассмотрены вопросы исследования вероятностных характеристик и построения прогнозов для выпадения осадков на основе специального преобразования исходных данных, позволяющего выделять паттерны в наблюдениях.

2. Исследование вероятностных характеристик осадков на основе паттернов

В качестве исходных данных рассматриваются объемы суточных осадков, наблюдавшихся в течение 60 лет в Потсдаме. Рассмотрим преобразование таких неотрицательных данных V_{daily} по следующему правилу: если в i -й день наблюдалась какая-либо положительная величина, то она заменяется на единицу ($\tilde{V}_{daily}^{(i)} = 1$), иначе значение $\tilde{V}_{daily}^{(i)}$ равно нулю. Таким образом, исходный ряд, состоящий из непрерывных значений, становится дискретным, принимающим два возможных значения $\{0, 1\}$. Данное упрощение позволяет анализировать непосредственно наличие или отсутствие осадков безотносительно к их объему. Таким образом, любая последовательность «сухих» (без осадков) и «дождливых» (с любым ненулевым объемом) дней представляет собой цепочку – паттерн – из «0» и «1».

Для каждого такого набора можно в рамках исторических данных определить частоты появления как отношение числа таких паттернов фиксированной длины N к общему числу возможных наборов (очевидно, 2^N), то есть фактически получить

значения вероятностей (согласно классическому определению). В рамках настоящего исследования были проанализированы наблюдения за 60 лет для значений параметра N от 1 до 14. Для каждого набора были получены значения частот (вероятностей), определен паттерн с максимальным значением [5].

В большинстве работ, посвященных статистическому анализу метеорологических данных, считается, что продолжительность периода выпадения осадков, измеренная в сутках (то есть число последовательных «дождливых» дней), подчиняется геометрическому распределению вероятностей. Возможно, данные предположения базируются на классической интерпретации геометрического распределения в терминах испытаний Бернулли как распределения числа последовательных «дождливых» дней («успех») до первого дня без осадков («неудача»). С использованием паттернов продемонстрировано [5], что последовательность «дождливых» и «сухих» дней не является даже марковской, поэтому использование схемы Бернулли, основанной на независимости испытаний, некорректно.

3. Построение прогнозов для осадков на основе паттернов

Используя данные о частотах (вероятностях) для паттернов, можно вычислять значения условных вероятностей появления в будущем определенных комбинаций, то есть получать вероятностные прогнозы для тех или иных событий. Например, при условии текущих наблюдений вида «1-1-0-0», то есть два дня подряд выпали осадки, в следующие двое суток они не регистрировались, можно формулировать утверждения следующего вида: «Вероятность осадков *через* 2 дня в Потсдаме при текущих наблюдениях составляет 0,3961, а вероятность отсутствия осадков *через* 2 дня – 0,6039». В отличие от стандартной для анализа данных практики, когда предсказываемое окно не должно превышать размер входных наблюдений, для исторических значений это правило может нарушаться.

В качестве альтернативного инструмента прогнозирования были использованы [5] нейронные сети прямого распространения с несколькими скрытыми слоями и различными функциями активации. В качестве обучающих рядов задействуются те же самые паттерны, о которых шла речь выше. Однако в явном виде частота каждого из наборов не используется, а соответствующие процедуры реализуются в скрытых слоях нейронной сети. В качестве результата работы получается прогноз на следующие за входным набором 1 – 2 дня. Наилучшей продемонстрированной точностью прогнозирования для нейронной сети с сигмоидной функцией активации и двумя скрытыми слоями была величина 82% для однодневного и 74% для двухдневного прогнозов (с реализацией на языке PHP). Введение дополнительного скрытого слоя, изменение функции активации на `rectifier`, увеличение объема входной выборки и использование библиотеки глубокого обучения `Keras` для языка `Python` позволили повысить точность прогнозов для тех же самых данных до 97% для однодневных и до 89% для двухдневных прогнозов.

4. Заключение

Для анализа вероятностного поведения процесса выпадения осадков и построения прогнозов предложено использование цепочек событий (паттернов), выделенных из данных. Продемонстрирована высокая точность прогнозов для нейронных сетей, при этом анализ основан исключительно на базовых статистических данных без привлечения каких-либо дополнительных сведений о метеорологических условиях.

Работа с паттернами представляет интерес с точки зрения верификации ансамблей прогнозов. Также данная методология может быть использована для прогнозирования поведения моментных характеристик конечных смесей вероятностных распределений [6] для определения направления изменения тренда в рамках моделирования физических экспериментов [7].

Благодарности

Работа выполнена при поддержке гранта РФФИ № 17-07-00851 и Стипендии Президента Российской Федерации молодым ученым и аспирантам (СП-538.2018.5). Автор выражает признательность член-корреспонденту РАН профессору С. К. Гулеву за предоставленные данные, профессору В. Ю. Королеву за полезные обсуждения в рамках совместных исследований метеорологических явлений и В. Кузьмину за реализацию обучения нейронных сетей.

Литература

1. Portmann R. W., Solomon S., Hegerl G. C., Spatial and seasonal patterns in climate change, temperatures, and precipitation across the United States // Proceedings of the National Academy of Sciences of the United States of America. — 2009. — Vol. 106 (18). — Pp. 7324–7329. doi:10.1073/pnas.0808533106.
2. Krzysztofowicz R. The case for probabilistic forecasting in hydrology // Journal of Hydrology. — 2001. — Vol. 249 (1–4). — Pp. 2–9. doi:10.1016/S0022-1694(01)00420-6.
3. Moustris K.P., Larissi I.K., Nastos P.T., Paliatsos A.G. Precipitation Forecast Using Artificial Neural Networks in Specific Regions of Greece // Water Resources Management. — 2011. — Vol. 25 (8). — Pp. 1979–1993. doi:10.1007/s11269-011-9790-5.
4. Ban N., Schmidli J., Schar C. Heavy precipitation in a changing climate: Does short-term summer precipitation increase faster? // Geophysical Research Letters. — 2015. — Vol. 42 (4). — Pp. 1165–1172. doi:10.1002/2014GL062588.
5. Горшенин А. К. Анализ вероятностно-статистических характеристик осадков на основе паттернов // Информатика и ее применения. — 2017. — Т. 11 (4). — С. 38–46. doi:10.14357/19922264170405.
6. Горшенин А. К. Концепция онлайн-комплекса для стохастического моделирования реальных процессов // Информатика и ее применения. — 2016. — Т. 10 (1). — С. 72–81. doi:10.14357/19922264160107.
7. Батанов Г. М., Горшенин А. К., Королев В. Ю., Малахов Д. В., Скворцова Н. Н. Эволюция вероятностных характеристик низкочастотной турбулентности плазмы в микроволновом поле // Математическое моделирование. — 2011. — Т. 23 (5). — С. 35–55.

UDC 004.032.26

On Investigation of Parameters of Meteorological Models Based on Patterns

A. K. Gorshenin

*Federal Research Center “Computer Science and Control” of RAS
44-2 Vavilova str., Moscow 119333, Russian Federation*

Email: agorshenin@frccsc.ru

The use of chains of events (patterns) extracted from data to analyze precipitation processes is suggested in the paper. The issues of testing the Markov property of data, probabilistic and neural network forecasting for statistical observations without involvement of any additional information about meteorological conditions are investigated. The best accuracy for neural networks trained on patterns is 97% for one-day and 89% for two-day forecasts.

Key words and phrases: precipitation, patterns, forecast, neural networks, deep learning, probabilistic forecasting.

UDC 004.4

Queueing Models for Description of Operation of Nodes of Wireless Sensor Networks with Energy Harvesting

A. N. Dudin*[†], S. A. Dudin*[†], O. S. Dudina*[†]

* *Department of Applied Mathematics and Computer Science
Belarusian State University
Independence Ave., 4, Minsk, 220030, Belarus*

† *Institute of Applied Mathematics and Telecommunications
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: dudin-alexander@mail.ru, dudin85@mail.ru, dudina_olga@email.co

Wireless sensor networks have a lot of important applications in various real-world systems. Nodes of sensor networks, which collect and transmit information about some objects, have small batteries with limited power and storage space. When the battery of a node is discharged, it should be charged again or replaced. Otherwise and the node terminates its operation and, eventually, the network may not be able to perform its designated task. Recent advances in developing various energy harvesting technologies have resulted in the design of new types of sensor nodes which are able to extract energy from the surrounding environment. The major sources of energy harvesting include solar, wind, sound, vibration, thermal, and electromagnetic power. The energy harvested from existing environmental sources is accumulated in energy storage devices. Design of such nodes requires, in particular, making decisions about the required equipment for energy harvesting and accumulation. The useful tool for decision support regarding the design of such nodes is provided by queueing theory. In this presentation, short survey of queueing literature devoted to optimization of operation of energy harvesting sensor nodes and its relation to the previously studied in the literature queueing models are presented.

Key words and phrases: wireless sensor network, energy harvesting, queueing analysis.

1. Introduction

The wireless sensor networks are widely used nowadays in a variety of applications, including safety, security, military purposes, monitoring of environment, transportation, agriculture, tracking of animals and human health. Nodes of sensor networks are usually provided by small batteries with limited power and capacity. Sometimes, the nodes are located in hard-to-reach places, so the problem how to charge the batteries is important. Recent advances in energy harvesting technology have resulted in the design of new types of sensor nodes which are able to extract energy from the surrounding environment, see, e.g. [1–5].

2. Analysis of wireless sensor networks by means of queueing theory

The simplest queueing model of operation of the sensor consists of a single server and two buffer. One buffer is designed for queueing the signals (customers) which have to be transmitted. The second buffer (storage) is designed for accumulation of energy. Capacity of the storage is finite. While energy is harvested by the node in continuous manner, following [6], it is useful to discretize it to the so called energy units. Transmission (service) of a signal requires availability of one or several energy units. If the required number of energy units is not available, service is postponed. When the service starts, immediately the corresponding number of energy units disappears.

If the arrival of customers and energy units is described by stationary Poisson processes and service time has the exponential distribution, the dynamics of the described queueing system is defined by two-dimensional Markov process

$$\xi_t = \{i_t, k_t\}, \quad t \geq 0, \quad i_t \geq 0, \quad k_t = \overline{0, N},$$

where i_t is the number of customers in the system, k_t is the number of energy units in the system at the moment t and N is capacity of the storage. Analysis of the process ξ_t is usually implemented via derivation of its ergodicity condition, computation of stationary distribution and basic performance measures of the system.

3. Related classes of queueing models and some references

The described queueing model is close to the queues with paired customers and assembly-like queues, see [7–10]. Another close class is given by queueing-inventory models, see [11–13]. Recent works devoted to analysis of queueing model of operation of the sensor node are [14–19].

4. Conclusions

Operation of a node of the wireless sensor network with energy harvesting is described by the queueing system and some relevant information about such type of queueing systems is provided.

Acknowledgments

This research has been prepared with the support by RUDN University Program 5-100.

References

1. H. H. R. Sherazi, L. A. Grieco, G. Boggia, A Comprehensive Review on Energy Harvesting MAC Protocols in WSNs: Challenges and Tradeoffs, *Ad Hoc Networks*, vol. 71, pp. 117–134, 2018.
2. W. Guo, Y. Deng, A. Nallanathan, B. Li, C., Zhao, C. . Multi-Scale Energy Harvesting, In *Wireless Information and Power Transfer: A New Paradigm for Green Communications*, 157–185). Springer, Cham, 2018.
3. V. Sharma, U. Mukherji, V. Joseph, and S. Gupta, Optimal energy management policies for energy harvesting sensor nodes, *IEEE Transactions on Wireless Communications*, vol. 9, no. 4, pp. 1326–1336, 2010.
4. K. Tutuncuoglu and A. Yener, Optimum transmission policies for battery limited energy harvesting nodes, *IEEE Transactions on Wireless Communications*, vol. 11, no. 3, pp. 1180–1189, 2012.
5. J. Yang and S. Ulukus, Optimal packet scheduling in an energy harvesting communication system, *IEEE Transactions on Communications*, vol. 60, no. 1, pp. 220–230, 2012.
6. E. Gelenbe, Synchronising energy harvesting and data packets in a wireless sensor, *Energies*, vol. 8, pp. 356–369, 2015.
7. G. Latouche, Queues with paired customers, *Journal of Applied Probability*, vo. 18(3), pp. 684–696, 1981.
8. J.M. Harrison, Assembly-like queues, *Journal of Applied Probability*, vol. 10(2), pp. 354–367, 1973.
9. U.N. Bhat, Finite capacity assembly-like queues, *Queueing Systems*, vol. 1(1), pp. 85–101, 1986.
10. E. Evdokimova, K. De Turck, D. Fiems, Coupled queues with customer impatience, *Performance Evaluation*, vol. 118, pp. 33–47, 2018.
11. M. Schwarz, C. Sauer, H. Daduna, R. Kulik, R. Szekli, *M/M/1 queueing systems with inventory*, *Queueing Systems*, vol. 54(1), pp. 55–78, 2006.
12. A. Krishnamoorthy, R. Manikandan, D. Shajin, Analysis of a multiserver queueing-inventory system, *Advances in Operations Research*, 2015.

13. A. Krishnamoorthy, B. Lakshmy, R. Manikandan, A survey on inventory models with positive service time, *Opsearch*, vol. 48(2), pp. 153–169, 2011.
14. K. Patil, K. De Turck, D. Fiems A two-queue model for optimising the value of information in energy-harvesting sensor networks, *Performance Evaluation*, 2018
15. S.A. Dudin, M.H. Lee, Analysis of single-server queue with phase-type service and energy harvesting. *Mathematical Problems in Engineering*, vol. 2016, ID592794, pp. 1–16, 2016.
16. A. Dudin, O. Dudina, Analysis of the MAP/PH/1 service system with repeat calls and energy audit, *Automatic Control and Computer Science*, vol. 45(5), pp. 277–285, 2015.
17. A. N. Dudin, M. H. Lee, S. A. Dudin, Optimization of Service Strategy in Queueing System with Energy Harvesting and Customers Impatience, *Applied Mathematics and Computer Science*, vol. 26, pp. 367–378, 2016.
18. C. S. Kim, A. Dudin, S. Dudin, O. Dudina, Performance Evaluation of a Wireless Sensor Node With Energy Harvesting and Varying Conditions of Operation, *Proceedings of IEEE International Communication Conference*, pp. 1-6, 2017.
19. A. Dudin, S. Dudin, O. Dudina, C. S. Kim. Analysis of a wireless sensor node with varying rates of energy harvesting and consumption, *Lecture Notes in Computer Science*, 10684, pp. 172–182, 2017.

УДК 004.4

Модели массового обслуживания для описания работы узлов беспроводных сенсорных сетей с самогенерацией энергии

А. Н. Дудин*[†], С. А. Дудин*[†], О. С. Дудина*[†]

* *Факультет прикладной математики и информатики
Беларуский государственный университет
пр. Независимости, д.4, Минск, Беларусь, 220030*

[†] *Институт прикладной математики и телекоммуникаций
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: dudin-alexander@mail.ru, dudin85@mail.ru, dudina_olga@email.co

Беспроводные сенсорные сети имеют множество важных приложений в различных реальных системах. Узлы сенсорных сетей, которые собирают и передают информацию о некоторых объектах, имеют небольшие батареи с ограниченными мощностью и пространством для хранения энергии. Когда аккумулятор узла разряжается, его необходимо снова зарядить или заменить. В противном случае, узел прекращает свою работу, и, в конечном итоге, сеть может не выполнить поставленные перед ней задачи. Недавние достижения в разработке различных технологий сбора энергии привели к созданию новых типов узлов сетей, которые способны извлекать энергию из окружающей среды. Основными источниками сбора энергии являются солнечная энергия, ветер, звуковая, вибрационная, тепловая и энергия электромагнитных волн. Энергия, собранная из существующих источников окружающей среды, накапливается в аккумуляторах энергии. Проектирование таких узлов требует, в частности, принятия решений относительно необходимого оборудования для сбора и накопления энергии. Полезный инструмент для поддержки принятия решений в отношении проектирования таких узлов обеспечивается теорией массового обслуживания. В этой презентации приведено краткое описание литературы о очередях, посвященной оптимизации работы узлов сбора энергии и их связи с ранее изученными в литературе моделями массового обслуживания.

Ключевые слова: беспроводные сенсорные сети, добыча энергии из окружающей среды, анализ очередей.

УДК 004.72

Исследование управления информационными и сетевыми ресурсами в мультисервисных сетях связи

Б. Г. Ибрагимов*, Р. Т. Гумбатов†, Р. Ф. Ибрагимов†

* Кафедра «Многоканальные телекоммуникационные системы»

Азербайджанский технический университет

проспект Г. Джавида, д. 25, Баку, Азербайджан, Az1073

† Лаборатория Теории телетрафика

Институт систем управления НАН Азербайджана
улица Б. Вахабзаде, д. 9, Баку, Азербайджан, Az1141

Email: i.bayram@mail.ru

Проанализированы современные ИКТ-технологии для создания единого инфокоммуникационного пространства на базе программно-конфигурируемых сетей, подсистема мультимедийной связи IMS и виртуализация сетевых функций. Исследованы управления информационными и сетевыми ресурсами в мультисервисных сетях связи общего пользования с использованием концепции единого инфокоммуникационного пространства и подучены аналитические выражения для оценки распределения и эффективного использования ресурсов пропускных способностей системы.

Ключевые слова: программно-конфигурируемые сети, единое инфокоммуникационное пространство, информационно-коммуникационные технологии, управление сетевыми ресурсами, виртуализация сетевых функций, IMS-архитектура.

1. Введение

В настоящее время с бурным развитием информационных и компьютерных технологий появилась возможность создания эффективных мультисервисных сетей связи на базе архитектурной концепции NGN (Next Generation Network), которые в рамках единой многооператорской инфраструктуры будут предоставлять пользователям мультимедийные услуги вторичных сетей общего пользования, сетей подвижных систем связи, пакетных, компьютерных сетей и IP-сетей, а также корпоративных сетей.

В соответствии с определением сетей NGN, пользующийся коммутацией пакетов обеспечивает предоставление основных, дополнительных и интеллектуальных услуг пользователям различных сетей с заданным качеством с использованием различных ИКТ-технологий. При этом уровень управления услугами содержит функции управления логикой услуг «Triple Play services» и приложений, которые представляет собой распределенную вычислительную и информационную среду [1, 2]. Последние в мультисервисных сетях связи на базе современных ИКТ-технологий требуют использование единой многооператорской инфраструктуры при помощи которого, обеспечиваются следующие: предоставление инфокоммуникационных услуг пользователям и управления услугами сети; создание и внедрение новых мультимедийных услуг; взаимодействия различных комплексных услуг [1, 3, 4].

Следует отметить, что данный подход требует для каждого оператора связи телекоммуникационных компаний определенных количествах абонентских и сетевых средств, коммутационных и транспортных систем, а также различных каналов связи и терминальных оборудований. Создания любого нового сетевого сервиса предполагает добавление всех новых наборов терминальных устройств, требующих места в аппаратных комнатах, новых источников питания. Всё это ведёт к росту стоимости оказываемых услуг системами связи и операционных затрат, а также приводит к тому, что затраты на развитие сети начинают опережать рост доходов, на которые направлены эти затраты, что не допустимо.

Учитывая вышеотмеченные, требуются новые подходы к созданию эффективных мультисервисных сетей связи, операторов связи телекоммуникационных компаний и

к развитию различных сервис-провайдеров сети связи, которые оказывают комплексные услуги пользователям с применением принципов совместного использования общих ресурсов.

Одним из таких подходов является построения распределенных мультисервисных сетей связи на базе архитектуры NGN, использующие современные ИКТ-технологии и концепции как программно-конфигурируемых сетей SDN (Software Defined Network), подсистема мультимедийной связи IMS (IP multimedia subsystem), так и виртуализация сетевых функций NFV (Network Functions Virtualization), что приобретает большую актуальность в Глобальном информационном обществе [3, 4].

Исследуемые архитектурные концепции программно-конфигурируемых сетей (ПКС), подсистема мультимедийной связи IMS и виртуализация сетевых функций (ВСФ) способствует построению новой и перспективной единой многооператорской платформы. Последние позволяют решить одну из главных телекоммуникационных задач – создание единого инфокоммуникационного пространства на базе современных ИКТ-технологий, обеспечивающие управления общими ресурсами в мультисервисных сетях общего пользования и управления услугами сети при оказании мультимедийных услуг с учетом независимо местоположения абонентов [4–6].

Наши работы посвящены решению задачи исследования эффективности управления информационными и сетевыми ресурсами в мультисервисных сетях связи общего пользования, базирующейся на концепции единого инфокоммуникационного пространства.

2. Общая постановка задачи

Системно-технические анализы показали [1–3], что эффективность внедрения функциональных компонентов и применение архитектурных концепций программно-конфигурируемых сетей SDN, подсистема мультимедийной связи IMS и виртуализация сетевых функций NFV в мультисервисных распределенных сетях связи открывают новые возможности использования информационных, канальных и сетевых ресурсов.

Создания единой многооператорской инфраструктуры и эффективность управления общими ресурсами в мультисервисных распределенных сетях связи позволяют предоставить пользователям современных и перспективных инфокоммуникационных услуг и приложений с применением принципов совместного использования.

Для создания многооператорской среды на базе концепции единой инфокоммуникационной инфраструктуры мультисервисных сетей связи состоит из трех важных аппаратных платформ и сегментов, которые дают возможности использования общих ресурсов и упрощают процедуру ввода дополнительных сервисов и интеллектуальных услуг для пользователей [2]: программно-конфигурируемые сети SDN, базируются на программно-аппаратных средствах и методах администрирования мультисервисных сетей общего пользования, а также управлять услугами сетей связи [3]. Основное свойство архитектуры SDN является отделение (абстрагирование) плоскости пересылки пакетов (Data plane) от плоскости управления (Control plane) при помощи стандартных протоколов между ними. Планирование сети и управление трафиком происходит программным путем. Для приложений верхнего уровня предоставляются интерфейсы прикладного программирования API. Таким образом, ввод новых услуг в сети ускоряется и облегчается. Узел коммутации ПКС сети в основном состоит из множества логических сетей, коммутатора и контроллера, которые взаимодействует протоколом OpenFlow.

ПКС сети является архитектурной концепцией для создания — сетей внутри сети с заранее определенными параметрами и конфигурацией. При этом сеть становится программируемой, создаваемой из имеющихся ресурсов под конкретные приложения и более открытой [3, 6].

- Виртуализация сетевых функций (ВСФ) – технология виртуализации физических сетевых элементов мультисервисные сети связи общего пользования, когда сетевые функции исполняются программными модулями, работающие

на стандартных серверах и виртуальных машинах в них. Эти программные модули могут взаимодействовать между собой для предоставления услуг связи, чем ранее занимались аппаратные платформы. NFV – это концепция виртуализации функций элементов сети оператора связи и её суть состоит в том, чтобы реализовать функции управления сетями и предоставления услуг программным путем, вместо того, чтобы использовать специализированное оборудование [1, 3–5].

Концепция ВСФ применима к любому оборудованию мультисервисных сетей связи общего пользования, включая маршрутизаторы, коммутаторы, межсетевые экраны и т.д. Ее основной особенностью является возможность оркестрации услуг, т.е. выделения виртуальных ресурсов тем или иным услугам по запросу. При этом достигается наиболее оптимальное использование информационных и сетевых ресурсов оборудования и серверов сети [6]. SDN и NFV, в общем, не зависят друг от друга, хотя NFV может в значительной степени дополнять SDN.

- Ядро-IMS является основой концепции мультисервисных сетей связи общего пользования и обеспечивает унифицированный доступ к приложениям независимо от технологии сети доступа, решает вопросы безопасности доступа со стороны терминальных устройств, эффективно управляет профилями абонентов. Стандарт IMS предполагает эволюционное развитие телекоммуникационной сети, обеспечивая при этом новые возможности для оптимизации архитектуры сети связи. В частности, при наличии надежной транспортной сети достаточно одного логического ядра IMS на всю сеть связи [1–3]. Функциональные компоненты IMS могут быть территориально-распределенными в зависимости от точек концентрации пользовательского трафика и вычислительных ресурсов ЦОД-ов. Такая возможность обеспечивает гибкий подход к созданию многооператорской среды на базе единой инфраструктуры сети связи. Виртуализация решения IMS дает возможность использования ресурсов «облачных» ЦОД-ов, независимость от аппаратной платформы, применение новых технологических механизмов резервирования и гибкого управления вычислительными и сетевыми ресурсами в зависимости от изменения нагрузки [3, 4, 6, 7].

В рамках доклада, в соответствии с постановкой задачи и обсуждением перечисленных вопросов, рассматривается управление информационными и сетевыми ресурсами в мультисервисных сетях связи с использованием многооператорской среды на базе технологий SDN/NFV и ядро IMS.

3. Описание и построение математической модели

В качестве эффективного управления информационными и сетевыми ресурсами в мультисервисных сетях связи общего пользования и управления услугами с использованием концепции единого инфокоммуникационного пространства выбраны распределение и эффективное использование канальных ресурсов и ресурсов пропускной способностей системы [2, 5].

С учетом управления канальными ресурсами в концепциях единого инфокоммуникационного пространства, сформулируем задачу распределения и эффективного использования ресурсов мультисервисных сетей связи общего пользования в наиболее общем виде, следуя [5]. При этом необходимо следует минимизировать среднюю задержку для одного типа потока пакетов и определить набор оптимальных пропускных способностей ПКС на базе коммутаторов и ядра-IMS при условии, что суммарная пропускная способность ограничена. Здесь, ограничения на суммарную пропускную способность означает, что заданы ограничения на суммарную стоимость оборудования сетей ПКС, платформы IMS и системы NFV.

Для решения выше поставленной задачи и учитывая важность взаимодействия программно-аппаратных средств, терминальных, сетевых и канальных оборудования сетей ПКС, платформы IMS и система NFV предлагается математические

модели, основанные на упрощениях в описании исследуемой концепции единого инфокоммуникационного пространства.

Математическая формулировка задачи предлагаемой математической модели для оценки характеристик пропускной способности и стоимости программно-аппаратных средств мультисервисных сетей связи описывается следующими целевыми функциями:

$$Q_{эфф} = \{ W [Arg \min_i (C_{i.ap})], i = \overline{1, n} \}, \quad (1)$$

при следующих ограничениях

$$N_k \leq N_{k.дон.}, T_{i.ср.з.} \leq T_{i.ср.з.дон.}, C_i \leq C_{i.маx доп.}, \eta_i \leq \eta_{i.дон.}, i = \overline{1, n} \quad (2)$$

где $C_{i.ap}$ — стоимость аппаратных и программных средств сетей ПКС, платформы IMS и система NFV, $i = \overline{1, n}$; C_i — пропускная способность подсистемы сетей ПКС, платформы IMS и система NFV при передаче i -го потока пакетов полезного и служебного трафиков, $i = \overline{1, n}$; $T_{i.ср.з.дон.}$ — среднее время задержки при передаче i -го потока пакетов, $i = \overline{1, n}$; η_i — коэффициент эффективного использования информационных и сетевых ресурсов единой многооператорской среда на базе концепции единой инфокоммуникационной инфраструктуры при передаче i -го потока пакетов, $i = \overline{1, n}$; N_k — в системе число используемых каналов; W — оператор совместной передачи полезного и служебного трафиков; $N_{k.дон.}$, $T_{i.ср.з.дон.}$, $C_{i.маx доп.}$, $\eta_{i.дон.}$ соответственно, допустимая значения величины пропускная способность подсистемы, среднее время задержки, число каналов и коэффициент эффективного использования информационных и сетевых ресурсов единой многооператорской среда при передаче i -го потока пакетов полезного и служебного трафиков, $i = \overline{1, n}$.

Выражения (1) и (2) определяют сущность рассматриваемой математической модели для оценки вероятностно-временных характеристик мультисервисных сетей связи на базе технологий построения распределенных сетей связи, использующие концепции единого инфокоммуникационного пространства при оказании мультимедийных услуг. Кроме того, (1) и (2) описывают особенности программно-аппаратных средств сетей ПКС, платформы IMS и технологий ВСФ, которые позволяют более точно учесть телекоммуникационные процессы, протекающие в исследуемых единой многооператорской инфраструктуре мультисервисных сетей связи общего пользования.

4. Оценка эффективности использования ресурсов

Допустим, что на буфер накопителя коммутатора многооператорской среды поступает стационарные гетерогенные пуассоновские потоки трафиков с параметрами λ_i , $i = \overline{1, n}$, создаваемые различными типами источников (голос, факс, Internet, данные, видео и др. в частном случае делим нагрузки на две большие группы — полезные и служебные нагрузки) нагрузки. На основе механизмов работы логических функций ядра IMS, сетей ПКС и ВСФ, а также особенностей обслуживания гетерогенных полезных $\lambda_{i.n}$ и служебных трафиков $\lambda_{i.c}$ разработана математическая модель (ММ) управления ресурсами в мультисервисных сетях связи.

Рассматриваемая ММ в общем случае представляет собой многоканальную систему массового обслуживания (СМО) общего типа $M/G/N_k/N_{бн} \leq \infty$ с очередями, учитывающие эффективность использования ресурсов единой многооператорской среды на базе технологий SDN/NFV и ядра IMS с использованием протокола OpenFlow, XMPP и SIP.

На основе ММ предполагаем, что гетерогенные пуассоновские потоки трафиков $\lambda_{i.n}$ и $\lambda_{i.c}$ обслуживаются программным коммутатором с пропускными способностями $C_n(\lambda_{i.n})$ и $C_c(\lambda_{i.c})$, и со скоростью обслуживаниями $\mu_n = C_n(\lambda_{i.n})/L_n^n$ и

$\mu_c = C_c(\lambda_{i.c})/L_n^c$, где L_n^n и L_n^c соответственно, средняя длина пакета полезного и служебного трафика.

Основываясь на формулы Хинчина–Полячека для системы $M/G/N_k/N_{bn} \leq \infty$ с очередями, определим среднее время задержки трафика:

$$E[T_{\text{ср.7}}] = \frac{\lambda}{N_k \cdot C} \cdot L_n \cdot \frac{\rho}{2\lambda(1-\rho)} + \lambda \cdot \frac{\sigma_{ts}^2}{2(1-\rho)} + L_n/(C \cdot N_k), \quad (3)$$

где C — суммарная пропускная способность и равно $C = \sum_{i=1}^2 C_i \cdot \max = C_n(\lambda_n) + C_c(\lambda.c)$; ρ — коэффициент загрузки всей сети ПКС/ВСФ и следующим образом находится: $\rho = \sum_{i=1}^2 (\lambda_i \cdot L_n/N_k \cdot C_i \cdot \max)$, $i = \overline{1, n}$.

Исследуемой ММ, основан на концепцию единого инфокоммуникационного пространства, сущность которое является разделение уровня управления сетью (системы и протоколы сигнализации NGN) и передачи полезных трафиков за счет переноса функций управления маршрутизаторами и коммутаторами в приложения, работающие на отдельном сервере контроллере.

На основе (3), можно минимизировать $E[T_{\text{ср.3}}]$, варьируя значениями $C_1(\lambda) = C_n(\lambda_n)$ и $C_2(\lambda) = C_c(\lambda.c)$ при фиксированной суммарной пропускной способности $C(\lambda) = C_1(\lambda) + C_2(\lambda)$, причем $C_1(\lambda) = C(\lambda) - C_2(\lambda)$. Учитывая последние, минимальные значения $E[T_{\text{ср.3}}]$ определяется при значении $C_{1,opt}(\lambda)$, которое, в свой очередь, находится из уравнения $dE[T_{\text{ср.3}}]/dC_{1,opt}(\lambda) = 0$.

С учетом вышеизложенного, можно определить коэффициент эффективного использования канального ресурса и ресурсы пропускной способности мультисервисных сетей связи общего пользования [5]:

$$\eta_{i.k}(\lambda) = (\lambda_i/C_{i,\max})(V_k \cdot T_S) \leq 1, \eta_{i.k}(C) = (L_n \cdot V_k/N_k \cdot C_{i,\max}) \leq 1, \quad (4)$$

где T_S — среднее время обслуживания трафика; V_k — скорость передачи канала связи.

Выполнения последнего условия позволяет определить резерв канального ресурса и ресурса пропускной способности сети связи:

$$R_{i.k} = 1 - \eta_{i.k}(\lambda), R_{i.nc} = 1 - \eta_{i.k}(C), \quad i = \overline{1, n}. \quad (5)$$

Таким образом, соотношения (4) и (5) позволяют более точно оценить эффективность использования и управления информационными и сетевыми ресурсами в мультисервисных сетях связи общего пользования, базирующейся на концепции единого многооператорского среда на базе технологий SDN/NFV и ядро IMS.

5. Заключение

В результате исследовании управления информационными и сетевыми ресурсами в мультисервисных сетях связи общего пользования с использованием концепции единого инфокоммуникационного пространства предложен ММ. На основе модели получены аналитические выражения, позволяющие оценить коэффициент эффективного использования канального ресурса и ресурса пропускной способности мультисервисных сетей связи общего пользования.

Литература

1. Ефимушкин В. А., Ледовских Т. В., Щербакова Е. Н. Инфокоммуникационное технологическое пространство цифровой экономику // Т-COMM: Телекоммуникации и транспорт.— 2017. — Т. 11, № 5. — С. 15–20.

2. Ибрагимов Б. Г., Гумбатов Р. Т., Ибрагимов Р. Ф. Исследование эффективности интеллектуальной сети связи при установлении соединения // *Материалы Всероссийской конференции с международным участием «Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем»*, РУДН, Москва, 2017. — С. 127–129.
3. Шалагинов А. В. SDN/NFV-технологическая основа цифровой трансформации // *Электросвязь*. — 2017. — № 12. — С. 88–89.
4. Волкова С. В. Использование современных ИКТ для создания единого инфокоммуникационного пространства // *Электросвязь*. — 2017. — № 12. — С. 98–99.
5. Ибрагимов Б. Г., Ибрагимов Г. Г. Исследование распределения ресурсов пропускной способности звена мультисервисных сетей связи // *Труды конференции «Телекоммуникационные и вычислительные системы»* Москва, МТУСИ, 2010. С. 67–68.
6. Логинов С. С. Об уровнях управления в программно конфигурируемой сети (SDN) // *T-Comm: Телекоммуникации и транспорт*. — 2017. — Т. 11, № 3. — С. 50–55.
7. Самуйлов К. Е., Чукарин А. В., Гайдамака Ю. В., Зарипова Э. Р. Математическая модель и метод оптимизации времени выполнения бизнес-процесса телекоммуникационной компании // *T-Comm: Телекоммуникации и транспорт*. — 2017. — Т. 11, № 1. — С. 37–44.

UDC 004.72

Research of Information Management and Network Resources in Multiservice Communication Networks

B. G. Ibrahimov*, R. T. Humbatov[†], R. F. Ibrahimov[†]

** Department of “Multichannel telecommunication systems”*

Azerbaijan Technical University

G.Dzhavid avenue, 25, Baku, Azerbaijan, Az10732

[†] Laboratory of Teletraffic Theory

Institute of Control Systems NAS Azerbaijan

B.Vahabzade, 9, Baku, Azerbaijan, Az1141

Email: i.bayram@mail.ru

The modern IKT-technologies have been analyzed to create a single infocommunication space on the basis Software defined networks, the IMS multimedia communication subsystem, and network functions virtualization. The management information and network resources in multiservice communication networks general use was studied using the concept of a single infocommunication space and analytical expressions were obtained for estimating the distribution and efficient use of the capacity resources of the system.

Key words and phrases: software defined networks, single infocommunication space, information and communication technologies, network resources management, network functions virtualization, IMS-architecture.

УДК 004.72

Исследование ресурсов в мультисервисных сетях связи на базе единого инфокоммуникационного пространства

Б. Г. Ибрагимов*, В. Р. Керимов†, Н. К. Гаджизаде†, А. М. Исаев†

* Кафедра «Многоканальных телекоммуникационных систем»
Азербайджанский технический университет
проспект Г. Джавида, 25, Баку, Азербайджан, Az1073

† Лаборатория теории телетрафика
Институт систем управления НАН Азербайджана
улица В. Вахабзаде, 9, Баку, Азербайджан, Az1141

Email: i.bayram@mail.ru

Задача исследования является одной из наиболее важных и сложных задач анализа эффективности управления информационными и сетевыми ресурсами в мультисервисных сетях общего пользования и управления услугами с использованием концепции единого инфокоммуникационного пространства. Целью данной работы является исследование постановки задачи управления ресурсами в сетях NGN. В работе рассматриваются построения математической модели эффективности управления ресурсами в мультисервисных сетях связи общего пользования и получены аналитические выражения для оценки распределения и эффективного использования ресурсов пропускных способностей системы.

Ключевые слова: система массового обслуживания, SDN, управления сетевыми ресурсами, FCAPS, виртуализация сетевых функций, оценка пропускной способности, IMS-архитектура, единые инфокоммуникационные пространства.

1. Введение

Интенсивное развитие телекоммуникационных и компьютерных сетей на базе архитектурной концепции NGN (Next Generation Network) с использованием современных телекоммуникационных технологий требует создания единого инфокоммуникационного пространства, которые способствуют выполнять задачи администрирование мультисервисных сетей связи общего пользования и управления услугами системы и сети [1].

На основе проведенных исследований установлено [1–6], что с целью эффективность управления информационными, вычислительными и сетевыми ресурсами в мультисервисных сетях связи целесообразно использовать единой многооператорской инфраструктуры и единого инфокоммуникационного пространства. Последние открывают новые возможности для оказания широкого спектра основных, дополнительных интеллектуальных услуг, а также «Triple Play services» благодаря совместным использованиям ресурсов единой инфраструктуры мультисервисных сетей связи общего пользования [7].

В настоящее время построения единого инфокоммуникационного пространства базируется на перспективных ИКТ-технологий и концепции подсистемы мультимедийной связи IMS (IP multimedia subsystem), программно-конфигурируемых сетей-SDN (Software Defined Network) и виртуализация сетевых функций-NFV (Network Functions Virtualization), обеспечивающие управления общими ресурсами в мультисервисных сетях связи общего пользования и управления услугами сети [5–8].

В данной работе рассматривается решение задачи — исследование и анализ эффективности управления ресурсами в мультисервисных сетях связи общего пользования и управления услугами с использованием концепции единого инфокоммуникационного пространства.

2. Описание и построение математической модели

Известно [3–6], что в сфере телекоммуникации увеличения набора количества мультимедийных услуг и объема передаваемых информационных и служебных трафиков вынуждает многооператорской инфраструктурой телекоммуникационных компаний увеличивать размер своей инфраструктуры, повышать эффективность управления ресурсами в мультисервисных сетях связи общего пользования, а также улучшить возможности использования имеющихся коммутаторов, многофункциональных терминальных оборудований и контроллеров системы.

Теперь, с учетом поступающих мультимедийных трафиков в ядро-IMS и звена коммутаторов сетей ПКС, создадим математическую модель (ММ) обслуживания потоков пакета трафиков. Данная система на основе алгоритмов работы функциональной архитектуры единого инфокоммуникационного пространства в общем случае представляет собой систему массового обслуживания (СМО) общего типа $M/G/N_k/N_{бн} \leq \infty$ с некоторыми допущениями при повышенной нагрузке $\rho_i \leq 1$, $i = \overline{1, n}$ (n – общее количество типов потока пакета трафика).

Рассматриваемая многоканальная система обслуживания имеет общий буферный накопитель с емкостью $N_{бн}$, где $N_{бн} \leq \infty$ и максимально доступным количеством обслуживающих коммутаторов системы N_k . На систему $M/G/N_k/N_{бн} \leq \infty$ с очередями поступающий сигнальный трафик на обслуживание является пуассоновским, с интенсивностью λ_i , $i = \overline{1, n}$ (здесь λ_i – это есть совокупность потоков пакетов информационных $\lambda_{i,u}$ и служебных $\lambda_{i,c}$ трафиков от различных источников сообщений, т.е. интенсивность потоков пакета трафика).

Для пуассоновских поступлений с интенсивностью λ_i , вероятность $P(v = k)$ поступления равна k -пакетов трафика за случайное время обслуживания τ . Допустим $f(\tau)$ является плотностью распределения τ и тогда выражается следующим образом [9]:

$$P(v = k) = \int_0^{\infty} P(v = k|\tau) \cdot f(\tau) d\tau, \quad (1)$$

где $P(v = k|\tau)$ – является вероятностью поступления k -пакетов узлов коммутатора сетей ПКС при условии фиксации конкретного значения τ и величину время τ можно интерпретировать как среднюю длину пакета $\tau = 1/\mu$ (при экспоненциальном времени обслуживания).

Следовательно, если есть очередь, то при наличии свободных ресурсов система, программным путем поступившие пакеты трафика начинает обслуживать с помощью новых узлов коммутатора. При этом в СМО длительность обслуживания i -го потоков пакетов служебных и полезных трафиков распределена произвольным законом и имеет функцию распределения $B(t) = P[B \leq t]$ с преобразованием Лапласа–Стилтьеса $b(s) = E[\exp(-bs)]$ и моментами b_i , $i = \overline{1, n}$. При этом условие является необходимым и достаточным для существования стационарного режима системы:

$$\rho = \sum_{i=1}^n (\lambda_i \cdot L_{i,n} / C_{i,\max} \cdot N_k) \leq 1, \quad i = \overline{1, n}, \quad (2)$$

где $L_{i,n}$ – длина передаваемого i -го пакета информационных $L_{i,u}$ и служебных $L_{i,c}$ трафиков и равно $L_{i,n} = L_{i,i} + L_{i,c}$; C_{\max} – максимальные значения пропускной способности узлов коммутатора системы при обслуживании i -го пакета трафика, $C_{\max} \leq 2, 50$ Гбит/с.

В данной системе ограничение $\rho_i < 1$, как и $\lambda_i \cdot \tau < 1$, $i = \overline{1, n}$, дает гарантию существования в СМО режима статистического равновесия.

Выражение (2) определяет коэффициент эффективного использования сетевого и канального ресурса с учетом возможности распределения пропускных способностей мультисервисных сетей связи общего пользования а использованием концепции единого инфокоммуникационного пространства.

3. Оценка показателей вероятностно-временных характеристик узлов коммутации сетей ПКС

На основе (1) можно определить порождающую функцию моментов для случайного значения v . Для пуассоновского распределения вида $P(v = k|\tau)$, производящую функцию $G_v(z)$ можно записать следующим выражением [9]:

$$G_v(z) = \sum_{k=0}^{\infty} P(v = k) \cdot z^k. \quad (2)$$

В (2) производя некоторые операции и изменяя порядок суммирования и интегрирования с учетом того, что под знаком суммы экспонента, получим выражение, которое представляет собой преобразование Лапласа плотности вероятности $f(\tau)$.

Учитывая вышепредложенные операции, выражение (2) примет следующий вид:

$$G_v(z) = F[\lambda(1 - z)], \quad (3)$$

Для конкретного случая функции $F(*)$ определяется, естественно, заданной плотностью $f(\tau)$.

Теперь, зная порождающую функцию $G_v(z)$, можно определить величину $E(v)$ следующим образом [9]:

$$E(v) = \left. \frac{dG_v(z)}{dz} \right|_{z=1} = \left. \frac{dF[\lambda(1 - z)]}{dz} \right|_{z=1} = \lambda \cdot E(\tau) = \lambda / (V_k \cdot N_k) = \rho \leq 1, \quad (4)$$

где $E(\tau)$ — первый момент распределения времени обслуживания и равен

$$E(\tau) = (L_n / C_{\max} \cdot N_k) = b_1, \quad b_1 = 1/\mu,$$

где μ^{-1} — среднее время обслуживания информационного и служебного трафиков.

Учитывая вышеизложенные и последние выражения (2), (3) и (4) можно определить среднее число пакетов полезных и служебных трафиков в буферном накопителе $E[L_n]$ в узле коммутатора сетей ПКС, следующим образом:

$$E[L_n] = [\sigma_v^2 + \rho(1 - \rho)] \cdot (2 - 2\rho)^{-1}, \quad (5)$$

где σ_v^2 — дисперсия число пакетов трафика, поступающих в течении времени обслуживания и с учетом коэффициента эффективного использования системы находятся следующим образом:

$$\sigma_v^2 = \sum_{k=0}^{\infty} (k - \rho)^2 \cdot P(v = k) = \frac{L_n}{C_{\max}} \cdot \lambda(1 + \lambda \cdot \sigma^2) \cdot N_k^{-1}, \quad (6)$$

где σ^2 — дисперсия распределения времени обслуживания потока пакетов трафика.

Одной из важных результатов исследования в данной системе $M/G/N_k/N_{\text{бн}} \ll \infty$ является средняя длина очереди в буферном накопителе и определяется следующим выражением:

$$E[L_{\text{cp}}] = \left[\lambda \cdot (1 + C_B^2) \cdot \frac{0,5E(\tau)}{1 - \rho} + 1 \right] \cdot \frac{L_n}{N_k \cdot C_{\max}} \cdot \lambda, \quad (7)$$

где C_B^2 — коэффициент вариация длительности обслуживания пакетов трафика.

Полученные выражения (5), (6) и (7) характеризуют показателем эффективности использования коллективного канального ресурса подсистемы мультимедийной связи IMS, сетей ПКС и виртуализация сетевых функций в мультисервисных сетях связи общего пользования, обеспечивающие управления общими ресурсами и управления услугами сети.

На основе модели и формулы Литтла определим среднее значения времени задержки обслуживания пакетов трафика и определяется следующим образом:

$$E[T_{\text{ср.з}}] = E[L_n] \cdot \left[\rho \cdot \frac{N_k}{L_n} \cdot C_{\max} \right]^{-1}.$$

Выражение (7) характеризует средним временем задержки сигнального трафика “End to end”, возникающая в процессе при передаче, при установлении соединения и при оказании мультимедийных услуг, а также является одним из основных параметров качества обслуживания — QoS пакетов трафика.

Таким образом, зная вероятностно-временные характеристики мультисервисных сетей связи с использованием концепции единого инфокоммуникационного пространства, сформулируем задачу распределения и эффективного использования ресурсов мультисервисных сетей связи в наиболее общем виде, следуя [9].

Для решения данной задачи было в качестве критерия оптимальности выбрана минимальная средняя задержка обслуживания трафика, возникающая при буферировании потоков пакетов трафика, поступающих от различных источников сообщения в систему. При этом необходимо определить оптимальное значения пропускных способностей МСС $C_{\max}^1, C_{\max}^2, \dots, C_{\max}^{(i)}, i = \overline{1, n}$ для каждого потока пакета трафика при условии, что суммарная пропускная способность ограничена некоторым значением $C_{\max} \text{ .огр.} \Rightarrow C_{\max}$.

Рассматриваемая данная задача с учетом коэффициента использования коллективного канального ресурса мультисервисной сети связи компактно описывается следующим образом:

$$\sum_{i=1}^n C_{\text{opt}}^{(i)}(\rho) = C_{\max}(\rho). \quad (8)$$

В (8) важным показателем является коэффициент загрузки системы при обслуживании потоков пакетов трафика.

Учитывая вышеизложенного, можно определить коэффициент эффективного использования ресурсы пропускной способности мультисервисных сетей связи общего пользования и управления услугами с использованием концепции единого инфокоммуникационного пространства следующим выражением [10]:

$$\eta_k(C_{\max}) = \sum_{i=1}^n (b_i \cdot \lambda_{i.\text{в.ых.}} / N_k) \leq 1.$$

Таким образом, в системе эта дополнительная пропускная способность получается благодаря распределению доли суммарной пропускной способности, оставшейся после выделения минимальных пропускных способностей для каждого полезного и служебного трафиков.

4. Заключение

В данной работе была рассмотрена задача исследования и анализ эффективности управления ресурсами в мультисервисных сетях связи общего пользования и управления услугами с использованием концепции единого инфокоммуникационного пространства.

Предложен ММ в виде СМО общего типа и получены аналитические выражения, которые позволяют оценить показателей эффективности управления ресурсами

в мультисервисных сетях связи общего пользования, а также выполнен анализ полученных результатов.

Литература

1. Ибрагимов Б. Г. Исследование и оценка ресурсов терминальных средств мультисервисных сетей телекоммуникаций // *Материалы Международная конференция «Проблемы функционирования информационных сетей»*. СО РАН, Новосибирск, 2008. — С. 56–60.
2. Степанов С. Н. Теория телетрафика: концепции, модели, приложения. — М.: Горячая линия – Телеком, 2015. — 868 с.
3. Логинов С. С. Об уровнях управления в программно конфигурируемой сети (SDN) // *T-Comm: Телекоммуникации и транспорт*. — 2017. — Т. 11, № 3. — С. 50–55.
4. Чукарин А. В. Об одном задаче управление сетевыми и информационными ресурсами в сетях NGN // *T-COMM*. — 2010. — № 7. — С. 123–125.
5. Волкова С. В. Использование современных ИКТ для создания единого инфокоммуникационного пространства // *Электросвязь*. — 2017. — № 12. — С. 98–99.
6. Stallings W. Software-defined networks and OpenFlow // *The Internet Protocol Journal*. — 2013. — Vol. 16, no. 1.
7. Ibrahimov B. G., Ismaylova S. R. The Effectiveness NGN/IMS Networks in the Establishment of a Multimedia Session // *American Journal of Networks and Communications*. — 2018. — Vol. 7, no. 1. — Pp. 1–5.
8. Chowdhury N. M. K. and Boutaba R. A Survey of Network Virtualization // *Computer Networks*. — 2010. — Vol. 54, no. 5. — Pp. 862–876.
9. Карташевский В. Г. Основы теории массового обслуживания. — М.: Горячая линия – Телеком, 2013. — 130 с.
10. Ибрагимов Б. Г., Ибрагимов Г. Г. Исследование распределения ресурсов пропускной способности звена мультисервисных сетей связи // *Труды конференции «Телекоммуникационные и вычислительные системы»*, Москва, МТУСИ, 2010.

UDC 004.72

The Investigation Resources in Multiservice Communication Networks on the Basis of the Single Infocommunication Space

B. G. Ibrahimov*, V. R. Karimov†, N. K. Hajzade†, A. M. Isayev†

* *Department of “Multichannel telecommunication systems”
Azerbaijan Technical University
G.Dzhavid avenue, 25, Baku, Azerbaijan, Az10732*

† *Laboratory of Teletraffic Theory
Institute of Control Systems NAS Azerbaijan
B.Vahabzade, 9, Baku, Azerbaijan, Az1141*

Email: i.bayram@mail.ru

The task of the research is one of the most important and complex tasks analyzing the effectiveness managing information and network resources in multiservice communication networks public use and service management using the concept of a single infocommunication space. The aim of this paper is to study the problem resource management in NGN networks. The paper considers the construction of a mathematical model resource management efficiency in multiservice communication networks general use and analytical expressions are obtained for estimating the distribution and efficient use system capacity resources

Key words and phrases: queuing system, SDN, network resource management, FCAPS, virtualization of network functions, bandwidth estimation, IMS-architecture, single infocommunication space.

УДК 519.872

Модель схемы распределения радиоресурсов в сети LTE в виде СМО с ненадежными приборами и конечной очередью

Е. А. Карнаухова, Д. В. Иванова, Е. В. Маркова

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: ek.karnauhova@gmail.com, daria.i1996@gmail.com, markova_ev@pfur.ru

Увеличение объема передаваемого трафика в мобильных сетях привело к необходимости поиска способов более эффективного управления радиоресурсами. Одним из возможных решений этой проблемы является система совместного использования лицензированного спектра частот LSA (Licensed Shared Access). В докладе предложена модель схемы сети LTE с системой LSA, представленная в виде системы массового обслуживания с ненадежными приборами и конечной очередью.

Ключевые слова: LTE, совместное использование радиочастот, LSA, СМО, ненадежные приборы, конечная очередь, показатели эффективности.

1. Введение

В работе рассматривается одна из возможных моделей схем распределения радиоресурсов при использовании системы LSA [1] в рамках беспроводной сети передачи данных LTE. Исследуется одна полоса частот — полоса совместного использования — частоты, взятые мобильным оператором в аренду у оператора-владельца. В случае отказа данной полосы, т.е. возврата ее оператору-владельцу, обслуживание пользователей не прерывается, а переходит в режим ожидания до момента восстановления полосы.

Далее перейдем к описанию модели в терминах теории массового обслуживания [2]. Будем рассматривать систему массового обслуживания (СМО) с ненадежными приборами, что соответствует полосе совместного использования, и конечной очередью.

2. Описание модели

Пусть СМО состоит из C ненадежных приборов, находящихся в доступном или недоступном состоянии, и очереди длины r , причем $r > C$. Предположим, что находящиеся на приборах заявки сохраняют за собой место в очереди. При этом приборы могут одновременно отказывать (переходить в недоступное состояние) с интенсивностью α и восстанавливаться (переходить в доступное состояние) с интенсивностью β . Входящий поток заявок имеет пуассоновское распределение с параметром λ . Среднее время обслуживания μ^{-1} распределено по экспоненциальному закону.

Пространство состояний системы имеет вид:

$$\mathbf{X} = \{0, 1, \dots, 2r + 1\},$$

где $n = 0, 2, \dots, 2k, \dots, 2r$ — состояния, в которых приборы недоступны, $n = 1, 3, \dots, 2k + 1, \dots, 2r + 1$ — состояния, в которых приборы доступны, а $k = 0, 1, \dots, r$ — число заявок в системе.

При поступлении новой заявки на обслуживание возможны следующие варианты:

- заявка поступает на обслуживание, занимает прибор и резервирует за собой место в очереди, если в системе имеются свободные приборы, находящиеся в доступном состоянии;

- заявка поступает в очередь, если в системе нет свободных приборов, находящихся в доступном состоянии, или приборы находятся в недоступном состоянии, а в очереди есть свободные места;
- заявка будет заблокирована, если в очереди нет свободных мест.

Диаграмму интенсивностей переходов для данной системы можно представить в виде, приведённом на рис. 1.

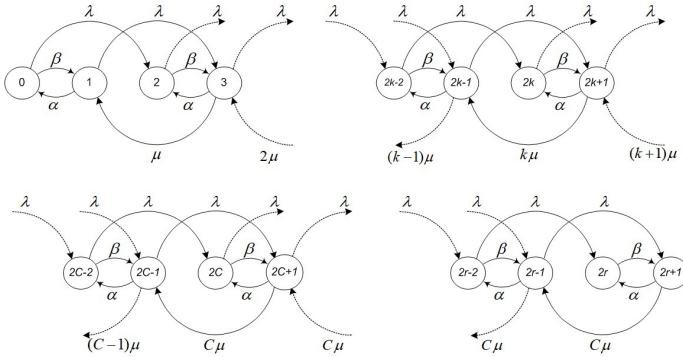


Рис. 1. Диаграмма интенсивностей переходов

С учетом сформулированных выше правил приема и обслуживания заявок элементы матрицы интенсивностей переходов определяются следующим образом:

$$a(n, n') = \begin{cases} \alpha, & \text{если } n' = n - 1, n = 2k + 1, k = 0, \dots, r; \\ \beta, & \text{если } n' = n + 1, n = 2k; \\ \lambda, & \text{если } n' = n + 2, n < 2r; \\ \frac{n-1}{2}\mu, & \text{если } n' = n - 2, n = 2k + 1, k = 1, \dots, C; \\ C\mu, & \text{если } n' = n - 2, n = 2k + 1, k = C + 1, \dots, r; \\ *, & \text{если } n' = n; \\ 0, & \text{в противном случае;} \end{cases}$$

где $*$ = $-(\frac{n-1}{2}\mu \cdot I(n = 2k + 1, k = 1, \dots, C) + C\mu \cdot I(n = 2k + 1, k = C + 1, \dots, r) + \alpha \cdot I(n = 2k + 1) + \beta \cdot I(n = 2k) + \lambda \cdot I(n < 2r))$.

В таком случае, систему уравнений глобального баланса можно представить в следующем виде:

$$\begin{aligned} & [\lambda \cdot I(n < 2r) + \frac{n-1}{2}\mu \cdot I(n = 2k + 1, k = 1, \dots, C) + \alpha \cdot I(n = 2k + 1) + \\ & + C\mu \cdot I(n = 2k + 1, k = C + 1, \dots, r) + \beta \cdot I(n = 2k)] \cdot p_n = \lambda \cdot I(n > 1)p_{n-2} + \\ & + \frac{n+1}{2}\mu \cdot I(n = 2k + 1, k = 0, \dots, C - 1)p_{n+2} + \alpha \cdot I(n = 2k)p_{n+1} + \\ & + C\mu \cdot I(n = 2k + 1, k = C, \dots, r)p_{n+2} + \beta \cdot I(n = 2k + 1)p_{n-1}, \quad k = 0, \dots, r, \end{aligned}$$

где p_n , $n \in \mathbf{X}$ — стационарное распределение вероятностей.

Основными характеристиками модели являются вероятность блокировки (1) и среднее число заявок в очереди (2)

$$B = \sum_{i=0}^1 p_{2r+i} = p_{2r} + p_{2r+1}; \quad (1)$$

$$Q = \sum_{k=1}^r k p_{2k} + \sum_{k=C+1}^r (k-C) p_{2k+1}. \quad (2)$$

3. Заключение

В ходе проделанной работы была рассмотрена модель одной LSA-полосы для нетолерантного (потокowego) к задержкам трафика. Исследование модели производилось для одномерного случая системы, была построена диаграмма интенсивностей переходов, выведены основные вероятностные характеристики системы — вероятность блокировки, средняя длина очереди всей системы, а также средняя длина очереди при доступных и недоступных приборах. Такое представление позволит в дальнейшем провести исследование зависимости рассмотренного случайного процесса от времени, что предоставит возможность максимально приблизить модель к реальным условиям.

Благодарности

Публикация подготовлена при финансовой поддержке РФФИ в рамках научных проектов № 18-07-00567 и № 18-37-00231.

Литература

1. Ponomarenko-Timofeev A., Pyattaev A., Andreev S., Koucheryavy Ye., Mueck M., Karls I. Highly dynamic spectrum management within licensed shared access regulatory framework // IEEE Commun. Mag. — 2016. — Vol. 54 (3). — Pp. 100–109.
2. Башарин Г. П. Лекции по математической теории телеграфика: Учеб. пособие. Изд. 3-е, испр. и доп. — М.: РУДН, 2009. — 342 с.

UDC 519.872

Resource Allocation Scheme Model in LTE Network as Queueing System with Unreliable Servers and Finite Queue

E. A. Karnauhova, D. V. Ivanova, E. V. Markova

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: ek.karnauhova@gmail.com, daria.i1996@gmail.com, markova_ev@pfur.ru

An increase in the volume of mobile traffic led to the need to find ways to control radio spectrum more effectively. One of the possible solutions to this problem is licensed shared access. In this report, we propose a scheme model of LSA framework in LTE network as queueing system with unreliable servers and finite queue.

Key words and phrases: LTE, binary use of spectrum, LSA, queueing system, unreliable servers, finite queue, performance measures.

УДК 621.39

Математическая модель звена доступа в сетях связи миллиметрового диапазона для одноадресных и многоадресных соединений

О. С. Крупко*, Ф. А. Москалева*, Р. Н. Ковальчуков*[†],
А. К. Самуйлов*[†], Д. А. Молчанов*[†]

* Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклушто-Маклая, д. 6, Москва, Россия, 117198

[†] Департамент электроники и телекоммуникаций
Технологический университет Тампере
ул. Коржеакоулункату, д. 1, Тампере, Финляндия, 33720

Email: alesiakrupko@yandex.ru, faina3586@mail.ru, rkovalchukov@gmail.com, samuylov_ak@rudn.university, molchanov_da@rudn.university

В данной работе предложена модель звена доступа в сетях связи миллиметрового диапазона при наличии трафика как многоадресных, так и одноадресных соединений. Модель является новой, поскольку размер запроса на ресурсы является случайной величиной учитывающей как случайное положение пользователя так и возможность нахождения в состоянии блокировки.

Ключевые слова: mmWave, мультимедиа, беспроводные сети.

1. Введение

В последнее время беспроводные сети миллиметрового диапазона частот (30 ГГц–300 ГГц) привлекли внимание научного сообщества благодаря возможности поддерживать высокую скорость передачи данных. Высокая скорость делает возможным использование таких сетей, как для предоставления современных широкополосных услуг связи, например, видео по запросу или транслирование нежатого видео, так и для услуг будущего поколения, например, услуг дополненной реальности.

Физические свойства волн миллиметрового диапазона частот несут в себе как достоинства, так и недостатки. Среди достоинств — высокая направленность. Поскольку передающие лучи генерируются в определенных направлениях, снижаются потери энергии, а проблемы конфиденциальности и безопасности стоят не так остро. Также снижается уровень интерференции, что позволяет повысить пропускную способность каналов. К недостаткам следует отнести высокие потери энергии при распространении в среде и возможность блокировки объектами из-за слабой дифракционной способности волн. Например, любое препятствие, как человек или предмет мебели, может сильно ухудшить связь.

Анализ литературы показывает, что предыдущие работы по сетям миллиметрового диапазона были сосредоточены на оптимизации передачи данных одноадресных соединений. Многоадресная передача на основе миллиметровых волн слабо исследована как в отечественной, так и в зарубежной литературе [1, 2]. В данной работе предлагается модель процесса обслуживания соединений услуг мультимедиа и одноадресных соединений точкой доступа сети миллиметрового диапазона.

2. Основная часть

Модель предполагает два типа соединений — один тип для одноадресных и один тип для многоадресных запросов. Входящие параметры модели — интенсивности потоков заявок для одноадресных и многоадресных соединений, интенсивности обслуживания, требование к ресурсам. Требования к ресурсам для одноадресных соединений является постоянной величиной; требования к ресурсам для заявок мультимедиа является случайной величиной. Закон распределения требований к

ресурсам для заявок мультивещания определяется с учетом случайного положения пользователей в зоне обслуживания точки доступа и возможной блокировки пути распространения сигнала. Время обслуживания заявок одноадресных и многоадресных соединений распределено по экспоненциальному закону. При этом, время обслуживания запросов многоадресных соединений определяется временем обслуживания первой поступившей заявки этого типа соединений.

В нашей работе рассматриваются такие характеристики модели, как вероятность блокировки для одноадресных соединений в случае недостатка свободных приборов, вероятность блокировки для многоадресных соединений в случае недостатка свободных приборов, распределение и среднее значение количества занятых ресурсов. Для разработки за основу взята модель из главы 4 [3].

3. Заключение

Мы описали модель, объединяющую процесс обслуживания многоадресных и одноадресных соединений точкой доступа миллиметрового диапазона. Мультивещание в совокупности с одноадресным трафиком позволяет эффективно экономить ресурсы сети и снижает вероятность блокировки для всех типов соединений.

Благодарности

Исследование выполнено при частичной финансовой поддержке РФФИ в рамках научных проектов № 18-07-000576, 17-07-00845.

Литература

1. Sundaresan K., Ramachandran K., Rangarajan S. Optimal Beam Scheduling for Multicasting in Wireless Networks, in *MobiCom '09 Proceedings of the 15th annual international conference on Mobile computing and networking*, pp. 205-216, 2009.
2. Kim W., Song T., Pack S. Rate Adaptation for Directional Multicast in IEEE 802.11ad Networks, in *IEEE International Conference on Consumer Electronics*, pp. 364-365, January 2012.
3. Наумов В. А., Самуйлов К. Е., Гайдамака Ю. В. Мультипликативные решения конечных цепей Маркова: монография. — М.: Изд-во РУДН, 2015. — 159 с.

UDC 621.39

The Mathematical Model of the mmWave Access Point for Two Types of Traffic

A. S. Krupko*, F. A. Moskaleva*, R. N. Kovalchukov*[†], A. K. Samuilov*[†],
D. A. Moltchanov*[†]

* Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation

[†] Department of Electronics and Communications
Tampere University of Technology
1 Korkeakoulunkatu, Tampere, 33720, Finland

Email: alesiakrupko@yandex.ru, faina3586@mail.ru, rkovalchukov@gmail.com, samuilov_ak@rudn.university, molchanov_da@rudn.university

We formulate a model representing the service process of the millimeter-wave access point simultaneously handling multicast and unicast types of sessions. The proposed model takes into account stochastic behavior of resource requests and specifics of multicast service.

Key words and phrases: mmWave, multicasting, wireless networks.

УДК 519.872

Асимптотический анализ второго порядка двухфазной гибридной СМО

А. А. Назаров, А. А. Анисимова

*Кафедра теории вероятностей и математической статистики
Национальный исследовательский Томский государственный университет
пр. Ленина, д. 36, Томск, Россия, 634050*

Email: siberienne94@yandex.ru, nazarov.tsu@gmail.com

Рассматривается двухфазная система массового обслуживания с простейшим входящим потоком. Продолжительность обслуживания на обеих фазах имеет экспоненциальное распределение. Первая фаза содержит один обслуживающий прибор и бункер с конечным числом мест для ожидания равным $N - 1$. Заявки, пришедшие на первую фазу в тот момент, когда на ней находится N заявок, не получают обслуживания и покидают систему. Вторая фаза содержит один прибор и орбиту, куда поступают заявки, заставшие прибор занятым. Находясь на орбите, заявки через случайные промежутки времени, имеющие экспоненциальное распределение, делают попытки вновь обратиться за обслуживанием. Из-за различия дисциплин ожидания рассматриваемая система была названа гибридной. Методом асимптотического анализа в условиях большой задержки заявок на орбите был найден второй момент распределения числа заявок на орбите.

Ключевые слова: система массового обслуживания, RQ-система, простейший поток.

1. Введение

Теория массового обслуживания зародилась в начале XX века как инструмент для решения практических задач, возникающих в реальных системах, предназначенных для удовлетворения массовых запросов на выполнение каких-либо услуг. В середине XX века все большую роль стали играть телекоммуникационные системы: компьютерные и телефонные сети, системы передачи данных, радио, телевидение, мобильная связь и другие. В отличие от классических систем массового обслуживания [2], для телекоммуникационных систем характерна ситуация, при которой заявка, заставшая обслуживающий прибор занятым, не встает в очередь, а уходит на орбиту, откуда через некоторые промежутки времени предпринимает попытки вновь обратиться за обслуживанием. Такие модели описываются в виде систем массового обслуживания с повторными вызовами или RQ-систем (Retrial queueing system), которые были подробно изучены Дж. Р. Арталехо [1].

В данной работе исследуется двухфазная СМО, которая содержит бункер с конечным числом мест для ожидания перед первой фазой и орбиту на второй фазе. Из-за различия дисциплин ожидания рассматриваемая система была названа гибридной.

2. Основная часть

Рассмотрим гибридную СМО, на вход которой поступает простейший поток с параметром λ . Каждая заявка последовательно проходит 2 фазы. Продолжительность обслуживания на обеих фазах имеет экспоненциальное распределение с параметрами μ_1 и μ_2 соответственно. Первая фаза содержит один обслуживающий прибор и конечный бункер с числом мест для ожидания $N - 1$. Вторая фаза содержит один прибор и орбиту, куда поступают заявки, заставляющие прибор занятым. Находясь на орбите, заявки через случайные промежутки времени, имеющие экспоненциальное распределение с параметром σ_1 , делают попытки вновь обратиться за обслуживанием.

Будем рассматривать двумерный случайный процесс $\{i_1(t), i_2(t)\}$, где $i_j(t)$, $j = \overline{1, 2}$ — число заявок на j -й фазе в момент времени t . Этот процесс не является марковским, поэтому введём дополнительную переменную:

$$k(t) = \begin{cases} 0, & \text{прибор на 2-ой фазе свободен,} \\ 1, & \text{прибор на 2-ой фазе занят.} \end{cases}$$

Пусть $P_k(i_1, i_2, t) = P\{i_1(t) = i_1, i_2(t) = i_2, k(t) = k\}$ — вероятность того, что в системе в момент времени t содержится i_1 и i_2 заявок на первой и второй фазах соответственно, при этом прибор на второй фазе находится в состоянии k . Будем рассматривать систему в стационарном режиме, при котором вероятности $P_k(i_1, i_2, t) = P_k(i_1, i_2)$.

Введём частичные характеристические функции:

$$H_k(i_1, u) = \sum_{i_2=0}^{\infty} e^{ju i_2} P_k(i_1, i_2).$$

Перейдём к векторно-матричному представлению. Для этого введём вектор:

$$H_k(u) = \begin{bmatrix} H_k(0, u) & H_k(1, u) & \dots & H_k(N, u) \end{bmatrix}$$

В [3] была получена асимптотика первого порядка в условии большой задержки заявок на орбите. Далее будет рассмотрена асимптотика второго порядка.

Теорема. Пусть $i_2(t)$ — число заявок на орбите в момент времени t . Тогда выполняется равенство:

$$\lim_{\sigma \rightarrow 0} M \exp \left\{ jw \sqrt{\sigma} \left(i_2(t) - \frac{\kappa_1}{\sigma} \right) \right\} = \exp \left\{ \frac{(jw)^2}{2} \kappa_2 \right\},$$

где параметр κ_2 определяется из совместного решения уравнений:

$$\frac{\mu_2}{2} R_1(\kappa_1) E + \frac{\mu_1}{2} RDE - \mu_2 f_1(\kappa_2) E + \mu_1 (f_0(\kappa_2) + f_1(\kappa_2)) DE = 0,$$

$$f(\kappa_2) = (v(\kappa_2) M^T) (MM^T)^{-1}.$$

Здесь вектор $R(x)$ и параметр κ_1 были получены в [3], а блочные матрица M и вектор $v(\kappa_2)$, а также матрицы A и D имеют вид:

$$M = \begin{pmatrix} \lambda D^T - A - \kappa_1 I & \vdots & \mu_1 D + \kappa_1 I & \vdots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ \mu_2 I & \vdots & \lambda D^T - A - \mu_2 I + \mu_1 D & \vdots & 1 \end{pmatrix},$$

$$v(\kappa_2) = \begin{pmatrix} \mu_2 R_1(\kappa_1) + \kappa_2 R_0(\kappa_1) & \vdots & -\mu_1 RD - \kappa_2 R_0(\kappa_1) & \vdots & 0 \end{pmatrix},$$

$$A = \begin{pmatrix} \lambda & 0 & \dots & 0 & 0 \\ 0 & \lambda + \mu_1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda + \mu_1 & 0 \\ 0 & 0 & \dots & 0 & \mu_1 \end{pmatrix},$$

$$D = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}.$$

При доказательстве теоремы использовались замены $\sigma = \varepsilon$, $u = \varepsilon w$. Поэтому окончательно характеристическая функция числа заявок на орбите имеет вид:

$$M e^{juiz} = \exp \left\{ ju \frac{\kappa_1}{\sigma} + \frac{(ju)^2 \kappa_2}{2 \sigma} \right\}.$$

Здесь κ_2/σ — второй момент числа заявок на орбите.

3. Заключение

В данной работе был проведён асимптотический анализ второго анализа двухфазной гибридной системы массового обслуживания с орбитой на второй фазе в условии большой задержки заявок на орбите.

Благодарности

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 16-31-00292 мол а.

Литература

1. Artalejo J. R., Gomez-Corral A. *Retrial Queueing Systems: A Computational Approach*. — Berlin: Springer, 2008.
2. Гнеденко Б. В., Коваленко И. Н. Введение в теорию массового обслуживания. 2-е изд., перераб. и доп. — М.: Наука. Гл. ред. физ.-мат. лит., 1987. — 336 с.
3. Назаров А. А. Асимптотический анализ первого порядка двухфазной СМО с конечной очередью и орбитой в условии большой задержки заявок на орбите / А. А. Назаров, А. А. Анисимова // Материалы молодежной научной школы. Россия, Москва, 23–27 октября 2017 г.; под общ. ред. К. Е. Самуйлова, Е. А. Кучерявого, А. Н. Дудина. — М.: РУДН, 2017. — С. 175–178.

UDC 519.872

Second-Order Asymptotic Analysis of a Two-Phase Hybrid Queueing System

A. A. Nazarov, A. A. Anisimova

*Department of Probability Theory and Mathematical Statistics
National Research Tomsk State University
Lenin Ave 36., Tomsk, 634050, Russian Federation*

Email: siberienne94@yandex.ru, nazarov.tsu@gmail.com

A retrial queue with two phases of essential service is considered. Customers arrive at the system according to a Poisson process. Customers are served incrementally on each phase during an exponentially distributed interval of time. The first phase is a finite capacity, single server queue. Customers arriving when the queue is full have to leave the system without service. After being served on the first phase, customers demand a service on the second phase. The second phase includes one server and one orbit, which customers join, if they found the server busy. Being in the orbit, customers make attempts to be served. After being served on the second phase customers leave the system.

Key words and phrases: queueing system, retrial queue, Poisson flow.

УДК 004.4

Об этапах интеллектуального анализа данных при построении и совершенствовании бизнес-процессов (Process Mining)

В. В. Никитина, Д. М. Хуракай

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: nikitinavalerie@mail.ru, homelinkin@mail.ru

Работа посвящена основным аспектам Process Mining. Рассмотрено обнаружение бизнес-процессов из простейшего журнала событий на основе альфа алгоритма. Определены необходимые шаги для построения модели процессов. Представлено два вида бизнес-процессов. Реализована проверка соответствия журнала событий, а также последующее улучшение модели.

Ключевые слова: Process Mining, бизнес-процессы, анализ журнала событий, альфа-алгоритм.

1. Введение

В настоящее время растет популярность науки о данных. Вездесущность данных событий в сочетании с технологическим процессом позволяет организациям диагностировать проблемы, основанные на фактах, а не на вымысле. Система фиксирует их в журналах событий, после чего благодаря интеллектуальной обработке данных выявляются бизнес-процессы компании, строится модель и принимаются меры по ее улучшению. Таким образом, задача актуальна и является одной из самых перспективных для изучения.

2. Основная часть

Существует три основные задачи интеллектуального анализа данных:

- обнаружение (восстановление модели бизнес-процесса при помощи анализа журнала информационной системы);
- проверка соответствия (проверка соответствия модели бизнес-процесса экземплярам бизнес-процесса, полученным из событийного журнала информационной системы);
- последующее улучшение (расширение) исходных процессов посредством получения информации из журнала [2].

Рассмотрим этап «Обнаружение» на основе простейшего примера журнала событий.

Дан журнал событий «Возврат билетов», в котором собраны данные за некоторый промежуток времени одной туристической компании. Всего в журнале 6 событий, все они имеют разный сценарий действий. Для работы с журналом событий можно применить альфа-алгоритм, основанная идея которого состоит в том, что задачи, которые следуют друг за другом в журнале, должны быть непосредственно связаны в модели процесса [1].

На первом этапе альфа-алгоритмы нам необходимо выделить все порядковые отношения. Под порядковыми отношениями подразумевается последовательность событий, следующая друг за другом. Для обозначения порядкового отношения используется символ «>», например, после событие a идет событие b , $a > b$.

На втором этапе выделяют причинную связь, она выполняется, если $a > b$, но не $a \succ b$. Это справедливо для всех пар, кроме (b, c) и (g, c) (см. табл. 2).

Таблица 1

Порядковые отношения журнала событий «Возврат билетов»

$a > b$	$b > c$	$c > d$	$d > f$	$c > g$	$g > d$
$d > e$	$a > g$	$g > c$	$k > c$	$d > k$	$k > b$
$k > g$	$a > c$	$c > b$	$b > d$		

Таблица 2

Причинные отношения журнала событий «Возврат билетов»

$a \rightarrow b$	$a \rightarrow c$	$a \rightarrow g$	$c \rightarrow d$	$d \rightarrow f$	$g \rightarrow d$
$d \rightarrow e$	$k \rightarrow c$	$d \rightarrow k$	$k \rightarrow b$	$k \rightarrow g$	$b \rightarrow d$

Также существуют потенциальные параллельные отношения справедливы для $b||c$ и $g||c$. Далее, используем журнал событий «Возврат билетов» и создаем таблицу порядковых отношений. Эти идентифицированные отношения относятся к задачам, которые непосредственно следуют друг за другом в журнале.

Таблица симметрична относительно диагонали, содержит три отношения: причинные отношения (\rightarrow или \leftarrow), параллельные отношения ($||$), отношения отсутствие наследия (решетка) (см. рис. 1).

	a	b	c	d	e	f	g	k
a	#	\rightarrow	\rightarrow	#	#	#	\rightarrow	#
b	\leftarrow	#	$ $	\rightarrow	#	#	#	\leftarrow
c	\leftarrow	$ $	#	\rightarrow	#	#	$ $	\leftarrow
d	#	\leftarrow	\leftarrow	#	\rightarrow	\rightarrow	\leftarrow	\rightarrow
e	#	#	#	\leftarrow	#	#	#	#
f	#	#	#	\leftarrow	#	#	#	#
g	\leftarrow	#	$ $	\rightarrow	#	#	#	\leftarrow
k	#	\rightarrow	\rightarrow	\leftarrow	#	#	\rightarrow	#

Рис. 1. Представление матрицы рабочих процессов «Возврат билетов»

На основе полученной таблицы строится BPMN диаграмма.

3. Заключение

На основе трех основных задач интеллектуального анализа данных можно строить модели процессов. Один из самых удобных и простых инструментов интеллектуального анализа на этапе «Обнаружение» является альфа-алгоритм. Его работа заключается в том, что задачи, которые непосредственно следуют друг за другом в журнале, должны быть непосредственно связаны в модели процесса.

Литература

1. Dumas M. Fundamentals of Business Process Management // ACM Computing Classification, 1998.
2. Wil A. D. Process Mining: Data science in 2-nd edition Netherlands: Springer ScienceBusiness Media, 2015.

UDC 004.4

About the Stages of Data Mining in Creation and Improving Business Processes

V. V. Nikitina, D. M. Khurakay

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: nikitinavalerie@mail.ru, homelinkin@mail.ru

The work is devoted to the main stages of the technological process. Considered the opening of business processes from the simplest event log based on the alpha algorithm. The necessary steps for building process models are determined. There are two types of representation of business processes. Implemented the verification of the correspondence of the event log, as well as the subsequent improvement of the model.

Key words and phrases: Process Mining, business processes, event log analysis, alpha-algorithm.

УДК 519.872, 519.217

Система с выбором кратчайшей очереди при поступлении заявок и максимальной очереди при обслуживании заявок

Т. С. Орбелян*, И. С. Зарядов*[†], Т. А. Милованова*

* Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198

[†] Институт проблем информатики
Федеральный исследовательский центр «Информатика и управление» РАН
ул. Вавилова, д. 44, к. 2, Москва, Россия, 119333

Email: 1032121703@rudn.university, zaryadov_is@rudn.university, milovanova_ta@rudn.university

Исследуется система массового обслуживания, в которой при поступлении выбирается очередь с наименьшим числом заявок, а при обслуживании — с наибольшим числом заявок.

Ключевые слова: система массового обслуживания, вероятностные характеристики, система поллинга, система выбора кратчайшей очереди.

1. Введение

В данной работе рассматривается система массового обслуживания, объединяющая в себе систему поллинга [1, 2] с выбором на обслуживание из очереди с наибольшим числом заявок и систему выбора кратчайшей очереди («join-the-shortest-queue») [3–6] для поступления.

Цели работы:

- построить математическую модель для случая, когда заявки из разных накопителей обслуживаются с разными интенсивностями (неоднородное обслуживание), описать совместное распределение числа заявок посредством системы уравнений равновесия;
- рассмотреть частный случай, когда интенсивности обслуживания заявок из разных накопителей совпадают (однородное обслуживание), получить в явном виде стационарное распределение числа заявок в системе, а также стационарное распределение числа заявок в каждом из накопителей, попытаться получить временные характеристики;
- перенести результаты частного случая на общий случай.

Начальные результаты данной работы были представлены в рамках международной молодежной научной школы «Молодежная научная школа по прикладной теории вероятностей и телекоммуникационным технологиям (АРТСТ–2017) 2nd International School on Applied Probability Theory and Communications Technologies (Aptct–2017)» [7].

2. Основная часть

Исследуется система $M|M(\mu_1, \mu_2)||1||r_1, r_2$ [8] (в нотации Башарина–Кендалла), которая состоит из одного обслуживающего прибора, время обслуживания на котором имеет экспоненциальное распределение, и двух накопителей емкости $r_1 \leq \infty$ и емкости $r_2 \leq \infty$, причем $r_1 \neq r_2$.

В рассматриваемую систему поступает пуассоновский поток заявок интенсивности λ , причем поступающая заявка становится в тот накопитель, текущая длина которого минимальна (реализация дисциплины выбора кратчайшей очереди («join-the-shortest-queue»)), если же текущие длины накопителей совпадают (в том числе, если накопители пусты, а прибор занят), то с вероятностью p выбирается первый

накопитель, с дополнительной вероятностью $q = 1 - p$ — второй. Если система пуста, то поступающая заявка сразу переходит на обслуживание.

Обслуживание происходит следующим образом. Прибор выбирает на обслуживающие заявку из того накопителя, очередь в котором максимальна. В случае равенства количества заявок в накопителях с вероятностью p_1 выбирается первый накопитель, с дополнительной вероятностью $q_1 = 1 - p_1$ — второй. Заявки из первого накопителя обслуживаются по с интенсивностью μ_1 , заявки из второго накопителя — с интенсивностью μ_2 . В общем случае данные интенсивности обслуживания различны ($\mu_1 \neq \mu_2$).

Функционирование системы можно описать трехмерным случайным процессом $\xi(t) = (\xi_1(t), \xi_2(t), \mu(t))$. Здесь $\xi_1(t)$ — число заявок в первом накопителе в момент времени t (от 0 до r_1), $\xi_2(t)$ — число заявок во втором накопителе в момент времени t (от 0 до r_2), $\mu(t)$ — состояние прибора (0 — прибор свободен (только если система пуста), 1 — обслуживается заявка из первого накопителя, 2 — обслуживается заявка из второго накопителя). Множество состояний данного случайного процесса можно представить в виде $\mathcal{X} = \{(0)\} \cup \{(i, j, k)\}$, где $i = \overline{0, r_1}$, $j = \overline{1, r_2}$, $k = 1, 2$. Состояние $\{(0)\}$ соответствует случаю, когда система пуста.

В рамках определенных выше механизмов поступления заявок и выбора на обслуживание для системы с накопителями неограниченной емкости возможны только те состояния, для которых $|i - j| \leq 1$, то есть, количество заявок в первом и втором накопителях может различаться не более чем на единицу. В данной ситуации возможно приведение матрицы интенсивностей переходов случайного процесса к трехдиагональному виду и применение известных матрично-аналитических методов [7].

В предположении о существовании стационарного режима, обозначим через $p_{i,j,k}$, $i = \overline{0, r_1}$, $j = \overline{1, r_2}$, $k = 1, 2$, стационарную вероятность того, что в первом накопителе находится ровно i заявок, во втором накопителе — ровно j заявок, и на приборе находится заявка из накопителя с номером k . Стационарная вероятность p_0 соответствует состоянию, когда система пуста.

В частном случае, когда заявки из разных накопителей обслуживаются с одинаковой интенсивностью ($\mu_1 = \mu_2$) множество состояний \mathcal{X} представимо в виде $\mathcal{X} = \{(0)\} \cup \{(i, j)\}$, где $i = \overline{0, r_1}$ — число заявок в первом накопителе, а $j = \overline{1, r_2}$ — число заявок во втором накопителе, прибор занят обслуживанием. Стационарные вероятности обозначаются через $p_{i,j}$ ($i = \overline{0, r_1}$, $j = \overline{1, r_2}$) для случая, когда прибор занят и в первом накопителе находится ровно i заявок, а во втором накопителе — ровно j заявок, и p_0 — вероятностью простоя системы.

Если через \tilde{p}_i ($i = \overline{0, r_1 + r_2 + 1}$) обозначить стационарную вероятность того, что в накопителях ровно i заявок:

$$\tilde{p}_{2i} = p_{i,i}, \quad i \geq 0, \quad \tilde{p}_{2i-1} = p_{i-1,i} + p_{i,i-1}, \quad i \geq 1,$$

то тогда распределение числа заявок в накопителях (и в системе) представимо в геометрической форме, как для системы $M|M|1|r$.

3. Заключение

В данной работе дано определение системы с выбором накопителя поступающей заявкой по принципу кратчайшей очереди и выбором накопителя с наибольшим числом заявок прибором для дальнейшего обслуживания. Представлена общая модель (неоднородное обслуживание заявок из разных накопителей), а также приведено стационарное распределение числа заявок в системе для частного случая, когда заявки из разных накопителей обслуживаются прибором с одинаковой интенсивностью.

Благодарности

Публикация подготовлена при поддержке Программы РУДН «5-100» и при финансовой поддержке РФФИ в рамках научных проектов № 18-07-00692, 16-07-00766.

Литература

1. Самуйлов К. Е., Абаев П. О., Гайдамака Ю. В., Гудкова И. А., Королькова А. В., Кулябов Д. С., Шукина О. Н. Мультисервисные сети связи, М.: РУДН, 2013.
2. Вишнеvский В. М., Семенова О. В. Системы поллинга: теория и применение в широкополосных беспроводных сетях, М.: Техносфера, 2007.
3. Suh K., Diot Ch., Kurose J., Massoulié L., Neumann Ch., Towsley D., Varvello M. Push-to-peer video-on-demand system: Design and evaluation // *IEEE Journal on Selected Areas in Communications*. — 2007. — Vol. 25 (9). — Pp. 1706–1716.
4. Bramson M. Stability of Join the Shortest Queue Networks // *Ann. Appl. Probab.* — 2011. — Vol. 21 (4). — Pp. 1568–1625.
5. Введенская Н. Д. Конфигурация перегруженных серверов при динамической маршрутизации // *Пробл. передачи информ.* — 2011. — Т. 47 (3). — С. 80–95.
6. Япо Г. С., Милованова Т. А., Зарядов И. С. Интервальные оценки характеристик системы с оптимальным выбором, Материалы Девятнадцатой международной научной конференции «Распределенные компьютерные и телекоммуникационные сети: управление, вычисление, связь» (DCCN-2016), 21–25 нояб. 2016 г. : в 3 т.; под общ. ред. В. М. Вишнеvского и К. Е. Самуйлова. Т. 3: Молодежная школа-семинар = Youth School-Seminar, С. 445–451, 2016.
7. Орбелян Т. С. Математическая модель системы с выбором при поступлении и обслуживании // Материалы международной молодежной научной школы «Молодежная научная школа по прикладной теории вероятностей и телекоммуникационным технологиям (АРТСТ–2017) = 2nd International School on Applied Probability Theory and Communications Technologies (Aptct–2017)». Россия, Москва, 23–27 октября 2017 г.; под общ. ред. К. Е. Самуйлова, Е. А. Кучерявого, А. Н. Дудина. — М.: РУДН, 2017. — С. 197–199.
8. Бочаров П. П., Печинкин А. В. Теория массового обслуживания. — М.: РУДН, 1995.

UDC 519.872, 519.217

The System with the Selection of the Shortest Queue for Arrival and the Queue with the Maximum of Tasks for Servicing

T. S. Orbelyan*, I. S. Zaryadov*[†], T. A. Milovanova*

* Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation

[†] Institute of Informatics Problems, FRC CSC RAS
44-2 Vavilova str., Moscow 119333, Russian Federation

Email: 1032121703@rudn.university, zaryadov_is@rudn.university, milovanova_ta@rudn.university

The queuing system, in which the queue with the lowest number of tasks is selected for arrival and the task for service is taken from the queue with the highest number of applications, is considered.

Key words and phrases: queuing system, polling system, join-the-shortest-queue mechanism, the-longest-queue-is-served mechanism.

УДК 519.872.5

Анализ модели процесса обработки разнотипных исков в страховой компании

Ш. Г. Погосян

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: shushanna.pogosian0101@yandex.ru

В данной работе рассматривается система массового обслуживания (СМО) замкнутого типа на примере модели процесса обработки страховых исков в случае, когда страховая компания (СК) заключает договоры страхования разного вида.

Ключевые слова: сеть массового обслуживания, страховая компания, договор страхования, страховщик, клиент, иск.

1. Введение

За последние 10 лет значительная часть населения страны стала проявлять высокое понимание первостепенной важности страхования из-за риска потерь как финансовых, так и производственных. За этот период отечественный страховой рынок значительно прогрессировал. В связи с этим наиболее весомыми становятся исследования методов математического моделирования в страховании, обеспечивающие надежную и динамичную деятельность страховой компании, так как именно экономико-математические вычисления в конечном счете определяют первичные параметры страховых договоров.

В настоящее время неподдельный интерес создают модели актуарной математики, то есть такой математики, которая взаимосвязана со страховым делом. Актуарная математика использует теорию вероятностей, а также другие основополагающие идеи для построения моделей страховых систем. В связи с необходимостью применения вероятностных методов в страховой математике, возникло новое направление для исследования процессов риска на основе теории массового обслуживания (ТМО). Далее рассмотрим модель [2] процесса обработки страховых исков в случае, когда страховая компания (СК) заключает договоры страхования $r - 1$ типов.

2. Постановка задачи

Рассматривается случай, когда СК заключает со страхователями договор $r - 1$ типов. В качестве модели процесса обработки разнотипных исков может быть использована замкнутая многолинейная сеть МО с пуассоновским потоком заявок. Заявки типа i , $i = \overline{1, r - 1}$ обслуживаются в m_i -линейной СМО S_i (система типа $M/M/m_i$), где происходит оценка исков, после чего переходят на обслуживание в центральную систему S_r (система типа $M/M/m_r$), где происходят выплаты. Далее для ожидания наступления страхового случая заявки направляются в $m_0 = K$ -линейную внешнюю среду S_0 (система типа $M/M/K/0$). Из внешней среды заявки типа i снова поступают на обслуживание в систему S_i , $i = \overline{1, r - 1}$ (рис. 1).

Используя принцип глобального баланса [1], была построена СУР (1):

$$p(\vec{k})\mu_0i + \sum_{i=1}^{r-1} \mu_i(k_i) + \mu_r(k_r) = \sum_{i=1}^{r-1} p(\vec{k} + \vec{e}_0 - \vec{e}_i)u(k_i - 1)\mu_0i + \sum_{i=1}^{r-1} p(\vec{k} + \vec{e}_i - \vec{e}_r)u(k_r - 1)\mu_i(k_i) + p(\vec{k} + \vec{e}_r - \vec{e}_0)u(k_0 - 1)\mu_r(k_r). \quad (1)$$

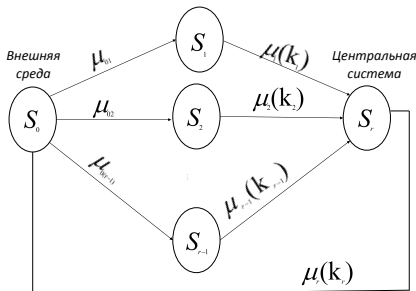


Рис. 1. Структура модели

3. Заключение

В данной статье была рассмотрена работа многолинейной СеМО с экспоненциальным входящим потоком заявок. Дисциплиной обслуживания заявок в системах сети является FIFO (First In, First Out). В данной работе исследованы модели процессов обработки страховых исков, позволяющие определить вероятностно-временные характеристики функционирования страховых компаний и решить ряд задач практического значения. Результаты данной работы могут быть использованы для анализа и оптимизации процессов функционирования страховых компаний.

UDC 519.872.5

The Analysis of the Model of Processing Different Types of Claims in an Insurance Company

Sh. G. Pogosian

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: shushanna.pogosian0101@yandex.ru

In the article, I consider the queueing network of closed type model on the example of model processing of insurance claims in case when the insurance company concludes insurance contracts of different type.

Key words and phrases: queueing network, insurance company, insurance contract, insurer, customer, claim.

Литература

1. Бочаров П. П., Печинкин А. В. Теория массового обслуживания: Учебник. — М.: Изд-во РУДН, 1995. — 529 с., ил.
2. Русилко Т. В., Маталыцкий М. А. Сетевые вероятностные модели обработки заявок в страховых компаниях. LAP Lambert Academic Publishing, 2012. — 327 с.

УДК 519.87

Применение информационных технологий для анализа показателей рейтинговой оценки вуза

О. Н. Ромашкова, Л. А. Пономарева, И. П. Василюк

*Кафедра прикладной информатики
ГАОУ ВО «Московский городской педагогический университет»
ул. Шереметьевская, д. 29, Москва, Россия*

Email: ox-rom@yandex.ru, ponomarevala@bk.ru, ipvkod@mail.ru

В работе построена модель прогнозирования рейтинговой оценки вуза на основе нейронной сети. Авторами описана методика моделирования, настройка параметров сети, приведены результаты моделирования.

Актуальность исследования: возможность применения прогнозной модели или методики ее построения для самостоятельного исследования.

Объектом исследования являются показатели деятельности российских университетов.

Предметом исследования является процесс прогнозирования рейтинга вуза.

Цель исследования — методологические аспекты построения нейросетевой модели прогнозирования рейтинговой оценки вуза с помощью инструментальных средств — пакета SPSS.

Ключевые слова: нейронные сети, многослойный перцептрон, моделирование, рейтинговая оценка вузов.

1. Введение

Для построения модели прогнозирования был использован большой массив данных с различной размерностью (таблица 1). В статистических методах обработки данных не имеет значения, каким способом будет минимизирована невязка целевой функции, модель будет оставаться неизменной. Авторами было принято решение анализировать показатели вузов с помощью нейронной сети. Сразу встает вопрос об архитектуре сети. Научная новизна исследования состоит в разработке методов и алгоритмов анализа и прогнозирования оценки деятельности вуза с применением нейронных сетей [1, 2].

Работа имеет практическую значимость, поскольку в ней проведена методика построения модели и настройка многослойного перцептрона в программе IBM SPSS Statistics.

2. Постановка задачи и экспериментальные данные

Исходные данные для моделирования представлены в таблице 1. Объектами исследования являются 1102 российских вуза. Свойства объектов – 123 показателя работы вузов. По приведенным показателям требуется предсказать значение целевой бинарной переменной – будет ли работа вуза эффективной. Используя программу IBM SPSS Statistics, необходимо построить нейронную сеть, которая входные данные разделит на кластеры и выявит их центры. Далее требуется с использованием обученной сети определить, к какому кластеру будет относиться новый входной вектор.

Таблица 1

Фрагмент экспериментальных данных

Ссылки	Наиме	Регион	Ведем	web-ca	Профи	Сведен	Наиме	Регион	Резуль	оценки	Ффекти	эсти деа	И1.1 (Ср	И1.2 (Ср	И1.3 (Ср	И1.4 (Ус	И1
1	http://ind.федерал.белгород.Министе.с	белгород	Министе.с	http://bse	образовательные организации сельскохозяйственной направленности	Значение			54,1	54,1	47,4	38,48					
2	http://ind.государст.белгород.Муницип.л	белгород	Муницип.л	http://bgi	образовательные организации творческой направленности	Значение			61,1	61,1	54,97	42,95					
3	http://ind.федерал.белгород.Министе.с	белгород	Министе.с	http://www	образовательные организации, не имеющие специфики деятельности	Значение			65,76	65,53	60,27	48,95					
4	http://ind.федерал.белгород.Министе.с	белгород	Министе.с	http://www	образовательные организации, не имеющие специфики деятельности	Значение			62,54	61,34	59,47	49,06					
5	http://ind.Автоном.белгород.Частные.с	белгород	Частные.с	http://www	образовательные организации, не имеющие специфики деятельности	Значение			0	0	48,46	39,3					
6	http://ind.Алексеев.белгород.Министе.с	белгород	Министе.с	http://af.1	образова: Сведения белгород.белгород.выполнил 4 или более показ	Значение			61,69	61,69	56,02	52,68					
7	http://ind.Старооск.белгород.Министе.с	белгород	Министе.с	http://www	образова: Сведения Национал.г.Москва	выполнил 4 или более показ	Значение		65,82	65,82	57,93	43,3					
8	http://ind.Старооск.белгород.Министе.с	белгород	Министе.с	http://www	образова: Сведения белгород.белгород.выполнил 4 или более показ	Значение			64,72	64,72	61,55	63,12					
9	http://ind.Старооск.белгород.Министе.с	белгород	Министе.с	http://www	образова: Сведения Российск.г.Москва	выполнил 4 или более показ	Значение		0	0	44,45	38,3					
10	http://ind.Филлал.А.белгород.Частные.с	белгород	Частные.с	http://www	образова: Сведения Воронеж Воронеж.выполнил 4 или более показ	Значение			0	0	47,95	38					
11	http://ind.федерал.брянская.Министе.с	брянская	Министе.с	http://www	образовательные организации, не имеющие специфики деятельности	Значение			54,38	54,38	52,59	39,46					
12	http://ind.федерал.брянская.Министе.с	брянская	Министе.с	http://www	образовательные организации, не имеющие специфики деятельности	Значение			59,34	59,34	54,01	46,15					
13	http://ind.федерал.брянская.Министе.с	брянская	Министе.с	http://www	образовательные организации, не имеющие специфики деятельности	Значение			64,42	64,42	60,75	46,81					
14	http://ind.Частное.о.брянская.Частные.с	брянская	Частные.с	http://www	образовательные организации, не имеющие специфики деятельности	Значение			0	0	45,8	33,1					
15	http://ind.федерал.брянская.Министе.с	брянская	Министе.с	http://www	образовательные организации, не имеющие специфики деятельности	Значение			51,73	51,73	46,63	40,28					
16	http://ind.брянский.брянская.Федерал.с	брянская	Федерал.с	http://www	образова: Сведения Московск.г.Москва	выполнил 4 или более показ	Значение		0	0	0	0					

3. Теоретическая часть

Для моделирования использовалась архитектура сети многослойный перцептрон. Этапы построения сети:

- оценить значимость показателей и определить диапазон изменения их значений;
- подготовить данные для моделирования;
- проектирование архитектуры сети — определение числа слоев и числа нейронов в каждом слое;
- обучение;
- тестирование.

4. Экспериментальные исследования

На этапе предварительной подготовки данных методом корреляционного анализа из 123 показателей для исследования оставлены 30.

При подготовке количественных переменных была учтена область определения и значения функции активации. Функция активации — сигмоида имеет область значения интервал (0, 1). В SPSS для приведения данных к интервалу (0, 1) использовалась нормализация. Область определения функции — вся числовая ось.

Число нейронов входного слоя сети равно числу независимых переменных — 30. Каждой зависимой переменной предназначается по одному выходному нейрону. Количество скрытых слоев SPSS определяет автоматически. Функция активации нейронов скрытого и выходного слоя — сигмоида.

Выборка была разделена на две части в пропорции: 60% — обучающая, 20% — данных в контрольной и проверочной. Разделение производилось программой случайным образом. Управление обучением происходило в мини-пакетном режиме, в котором алгоритм обратного распространения ошибки — это стохастический градиентный спуск. Правило остановки обучения сети: максимальное количество шагов без изменения ошибки.

5. Полученные результаты

Поскольку число скрытых слоев и количества нейронов в этих слоях программа выбирала автоматически, было построено две модели с разной архитектурой сети (таблица 2).

Таблица 2

Нейросетевые модели с разной архитектурой сети

Размер обучающей выборки	Архитектура сети		Процент ошибочных прогнозов	
	Скрытые слои	Число нейронов	Обучающая	Проверочная
441 (60%)	1	10	18,5%	17,9%
441 (60%)	2	200	18%	18%

Вычисления показали, что число слоев и нейронов не сильно влияют на качество модели. В результате исследования выборка была разделена на два кластера (табл. 3).

Таблица 3

Результаты классификации

Выборка	Предсказанные		Процент правильных
	1 кластер	2 кластер	
Обучающая	56,9%	43,1%	81,7%
Контрольная	56,8%	43,2%	82,1%
Проверочная	56,2%	43,8%	82,2%

Показатель площади под AUC (Area Under Curve) кривой равен 0,97, что говорит о хорошем качестве построенной модели (рис. 1).

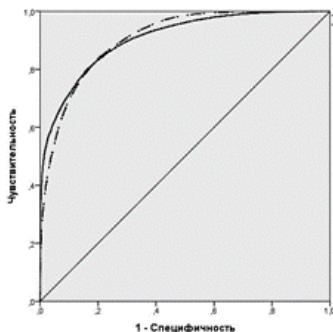


Рис. 1. AUC – кривая по построенной модели

6. Заключение

Авторами рассмотрена методика моделирования рейтинговой оценки вузов на примере построения нейронной сети в программе IBM SPSS Statistics. Данная методика может быть альтернативой статистическим методам исследования подобных экспериментальных данных [3, 4].

Литература

1. Пономарева Л. А., Коданев В. Л. Разработка модуля корпоративной информационной системы «Образовательная среда вуза» на базе облачных технологий // В сборнике: Информатика: проблемы, методология, технологии сборник материалов XVII международной научно-методической конференции: 5 т., 2017. — С. 393–398.
2. Ромашкова О. А., Моргунов А. И. Информационная система для оценки результатов деятельности общеобразовательных организаций г. Москвы // Вестник Российского университета дружбы народов. Серия: Информатизация образования. — 2015. — № 3. — С. 88–95.
3. Пономарева Л. А., Голосов П. Е. Разработка математической модели учебного процесса в вузе для повышения качества образования // Фундаментальные исследования. — 2017. — № 2. — С. 77–81.
4. Ромашкова О. Н., Ермакова Т. Н. Мониторинг качества образования в средней общеобразовательной организации с использованием современных средств информатизации // Вестник Российского университета дружбы народов. Серия: Информатизация образования. — 2014. — № 4. — С. 10–17.

UDC 519.87

Application of Information Technology for the Analysis of the Rating of the University

O. N. Romashkova, L. A. Ponomareva, I. P. Vasilyuk

*The Department of Applied Informatics at the Moscow City University
Sheremetevskaya str. 29, Moscow, Russia*

Email: ox-rom@yandex.ru, ponomarevala@bk.ru, ipvkod@mail.ru

A model for forecasting the rating of a university based on a neural network is being constructed. The authors describe the modeling technique, configure the network parameters, and present the simulation results.

The relevance of the study: the possibility of using the model itself or the methodology of its construction for independent research.

The object of research are the performance indicators of Russian universities.

The subject of the study is the process of forecasting the rating of the university.

The aim of the study is to build a model for forecasting the rating of the institution.

Key words and phrases: neural networks, multi-layer perceptron, modeling, rating of universities.

УДК 004.7:519.872

Адаптивные дисциплины опроса в системах поллинга и их имитационное моделирование

О. В. Семёнова*, Д. Т. Буй†

* *Институт проблем управления им. В.А. Трапезникова РАН
ул. Профсоюзная, д. 65, Москва, Россия, 117997*

† *Московский физико-технический институт
Институтский пер., д. 9, Долгопрудный, Московская обл., Россия, 141701*

Email: olgasmnv@gmail.com, duytan@phystech.edu

В статье рассматривается система поллинга с циклическим адаптивным порядком опроса очередей: адаптивным динамическим опросом и адаптивным упорядоченным опросом. При адаптивном динамическом опросе обслуживающий прибор (сервер) циклически посещает очереди, но при этом в цикле пропускает те из них, которые были пусты при их опросе в предыдущем цикле. При адаптивном упорядоченном опросе сервер в начале каждого цикла упорядочивает очереди в порядке убывания их ранга и в этом порядке их посещает. В следующем цикле процедура повторяется заново. Ранг — это число, которое присваивается каждой очереди в зависимости от того, сколько раз подряд она не была пуста в момент ее опроса. Если очередь опрошена пустой, то ранг ее снижается на единицу. Ставится задача имитационного моделирования таких систем и проводится сравнительный анализ разных типов адаптивного опроса.

Ключевые слова: системы поллинга, адаптивные дисциплины опроса, имитационное моделирование.

1. Введение

Системы поллинга представляют собой системы массового обслуживания с несколькими очередями (или несколькими потоками заявок). Обслуживающий прибор по определенному правилу посещает очереди и обслуживает находящиеся в них заявки. Развитие теоретических результатов анализа систем поллинга до 1995 г. изложено в монографии [1], а работы, опубликованные в 1996-2009 годах, отражены в обзорах [2, 3]. Обобщению и систематизации моделей и методов исследования стохастических систем с циклическим опросом и их применению для проектирования широкополосных беспроводных сетей посвящены монографии [4, 5].

Очереди системы поллинга обслуживаются согласно заданной дисциплине обслуживания. Она характеризуется числом заявок, которое может обслужить сервер за одно посещение очереди. В данной работе будем рассматривать *исчерпывающую дисциплину* (сервер обслуживает заявки до тех пор, пока очередь не опустеет) и *шлюзовую дисциплину* (сервер обслуживает лишь те заявки, которые находились в очереди в момент ее опроса сервером). Правило, следуя которому сервер выбирает очередь для обслуживания, называется порядком опроса. Наиболее распространенным порядком опроса является циклический опрос, когда сервер посещает очереди от первой до последней и вновь возвращается к первой очереди. В данной работе рассматривается адаптивный динамический порядок опроса очередей, а также предлагается адаптивный упорядоченный опрос, описание которого приведено далее.

2. Циклический адаптивный опрос

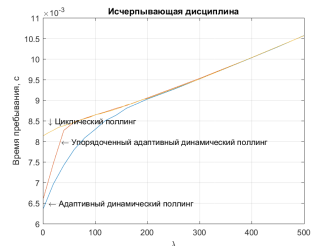
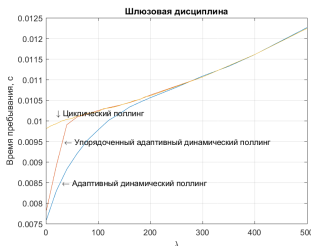
Адаптивный динамический опрос очередей введен в работе [6], при котором обслуживающий прибор циклически посещает очереди, но при этом в цикле пропускает те из них, которые были пусты при их опросе в предыдущем цикле. Анализ такой дисциплины, или порядка опроса, требует информации о состояниях очередей в предыдущем цикле опроса, что значительно его усложняет, и на данный момент

удалось построить лишь приближенный алгоритм расчета характеристик систем с адаптивным опросом.

Разновидностью адаптивного опроса можно считать адаптивный упорядоченный опрос, который описывается следующим образом. В первом цикле всем очередям присваивается число (ранг), равное нулю, а сервер опрашивает очереди в порядке их нумерации. Если в этом же цикле при подключении сервера к некоторой очереди она пуста, ранг этой очереди уменьшается на 1, в противном случае ранг увеличивается на 1. В следующем цикле сервер посещает очереди в порядке убывания их ранга, а процедура изменения значений рангов повторяется. Если несколько очередей имеют одинаковый ранг, сервер их обслуживает по порядку возрастания их номеров. Предполагаем, что ранг очереди может увеличиваться неограниченно, а при -1 уменьшаться уже не может.

3. Имитационное моделирование систем поллинга с адаптивным опросом

В данном разделе кратко представим результаты имитационного моделирования системы поллинга с двумя видами адаптивного опроса. Имитационная модель реализована с помощью OMNeT++ Discrete Event Simulator. Система поллинга состоит из сервера и 6 очередей. Потоки заявок в систему простейшие, времена обслуживания и переключения сервера между очередями распределены экспоненциально с параметрами, соответственно, 10000 и 1000 и не зависят от номера очереди. Для иллюстрации результатов предполагаем, что первые три очереди системы имеют идентичные параметры входного потока ($\lambda_1 = \dots = \lambda_3 = 1000$), а параметры следующих трех очередей также идентичны, но изменяются от 1 до 1500 (при этом загрузка всей системы меняется от 0,3 до 0,75). Будем рассматривать зависимость среднего взвешенного времени ожидания в системе W при обычном циклическом опросе очередей и при двух видах адаптивного опроса. Результаты расчетов показывают, что при исчерпывающем обслуживании очередей величина W всегда ниже, чем при шлюзовом. При изменении загрузки системы доля пропущенных циклов, в которых очередь (одна из очередей 4-6) не опрашивалась сервером, меняется от 0,49 до 0. Заметим, что нулевая доля пропущенных циклов вполне объясняется определением адаптивного опроса. На рисунках приведен сравнительный анализ различных видов опроса при изменении загрузки системы с точки зрения среднего взвешенного времени ожидания заявок во всех очередях. Заметим, что в данном примере при использовании шлюзовой дисциплины обслуживания выигрывает от использования адаптивного динамического опроса по сравнению с циклическим опросом достигает 29,9% (при $\lambda = 1$). При использовании адаптивного упорядоченного опроса этот выигрыш составляет 28,9%. При исчерпывающем обслуживании очередей эти проценты составляют, соответственно, 24,4% и 23,0%. Заметим, что если рассматривать группы очередей (с номерами 1-3 и 4-6) по отдельности, то адаптивные дисциплины позволяют снизить среднее время ожидания в каждой из очередей первой группы и, наоборот, они невыгодны для очередей второй группы с относительно низкой загрузкой.



4. Заключение

В работе описаны дисциплины адаптивного опроса в системах поллинга, построена имитационная модель таких систем и в последующих исследованиях будет решаться вопрос о целесообразности использования таких дисциплин опроса при различных критериях качества работы системы (например, в случае, если взимается плата за подключение сервера к очередям или за ожидание заявок в очередях).

Благодарности

Работа выполнена при финансовой поддержке Российского научного фонда и DST (Индия) (грант № 16-49-02021) в рамках совместного научно-исследовательского проекта Института проблем управления РАН и CMS Колледжа Коттаям.

Литература

1. Borst S. C. Polling systems. Amsterdam: Stichting Mathematisch Centrum, 1996.
2. Вишнеvский В. М., Семёнова О. В. Математические методы исследования систем поллинга // Автоматика и телемеханика. 2006. — № 2. — С. 3–56.
3. Вишнеvский В. М., Мишкой Г. К., Семёнова О. В. Новые модели и методы исследования систем поллинга // Proceedings of the International Conference proceedings Distributed Computer and Communication Networks. Theory and Applications (DCCN-2009, Moscow). — М.: R&D Company «Information and Networking Technologies», 2009. — С. 79–85.
4. Вишнеvский В. М., Семёнова О. В. Системы поллинга: теория и применение в широкополосных беспроводных сетях. — М.: Техносфера, 2007. — 309 с.
5. Vishnevsky V., Semenova O. Polling Systems: Theory and Applications for Broadband Wireless Networks // LAMBERT Academic Publishing, 2012. — 317 p.
6. Vishnevsky V. M., Dudin A. N., Klimenok V. I., Semenova O. V., Shpilev S. Approximate method to study $M/G/1$ -type polling system with adaptive polling mechanism // Quality Technology and Quantitative Management. — 2012. — Vol. 2. — Pp. 211–228.

UDC 004.7:519.872

Adaptive Polling Disciplines and Their Simulation Models

O. V. Semenova*, D. T. Bui†

* *Institute of Control Sciences of Russian Academy of Sciences
Profsoyuznaya str. 65, Moscow, 117997, Russia*

† *Moscow Institute of Physics and Technology
9 Institutskiy per., Dolgoprudny, Moscow Region, 141701, Russia*

Email: olgasmnv@gmail.com, duytan@phystech.edu

The paper deals with a polling system with cyclic adaptive polling orders: an adaptive dynamic polling and an adaptive ordered polling. With adaptive dynamic polling, the server polls the queues in a cyclic way but skips (does not visit) those that were empty when polling them in the previous cycle. With an adaptive ordered polling, at the beginning of each cycle the server orders the queues in the descending order of their rank and then starts polling them according to the prescribed order. In the next cycle, the procedure is repeated anew. A rank is a number assigned to each queue depending on how many cycles before it was polled non-empty consecutively. If the queue is polled empty then its rank is reduced by one. We propose a simulation model to compare different types of adaptive polling orders.

Key words and phrases: polling systems, adaptive polling order, simulation.

УДК 621.39

Анализ вероятности успешного установления соединения по радиоканалу случайного доступа с использованием виртуальных преамбул

И. Э. Синецны

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: iesinitsyn@gmail.com

В настоящее время стремительно растет число устройств, обменивающихся между собой информацией в сети LTE (Long-Term Evolution). Устройства МТС (Machine-type communications) эволюционируют в так называемые «low-cost МТС» (LC-MTC), предлагаемые консорциумом 3GPP (3rd Generation Partnership Project). Предлагаемые LC-MTC должны эффективно справляться с интенсивным доступом через множество узкополосных каналов случайного доступа (Random Access Channel, RACH). С увеличением количества и плотности размещения МТС-устройств схема произвольного доступа для LC-MTC RACH должна быть усовершенствована. В текущей работе предлагается использовать концепт виртуальных преамбул в процедуре установления соединения по радиоканалу случайного доступа, исследована зависимость вероятности успешного установления соединения от числа устройств МТС в соте.

Ключевые слова: канал случайного доступа, коллизия, вероятность успешного установления соединения, low-cost МТС, преамбула.

1. Введение

Межмашинные коммуникации крайне важны для эффективной передачи данных от устройств в сеть связи для обеспечения различных сервисов Интернета Вещей, например, таких как приложения для умного дома, логистического отслеживания, здравоохранения, безопасности, наблюдения и умного измерения [1–5]. Ожидается, что более 50 миллиардов устройств будут подключены к сети для обслуживания устройств Интернета Вещей и спрос на эффективные системы МТС-коммуникаций значительно увеличится [2].

Для снижения стоимости производства LC-MTC оборудования рабочая группа 3GPP RAN предложила новую структуру RACH, где сеть LTE присваивает несколько узкополосных каналов случайного доступа (NarrowBand-RACH, NB-RACH) в рамках своей пропускной способности. Устройство LC-MTC выбирает один из доступных каналов NB-RACH и устанавливает соединение, выполняя процедуру установления соединения по выбранному каналу RACH. В отличие от классической процедуры, где устройство МТС не выбирает физический канал, в новой структуре с дополнительным идентификатором выбор конкретного канала может уменьшить коллизии, возникающие при инициации установления соединения. Таким образом, введение дополнительных идентификаторов позволит уменьшить потребляемое количество энергии [6]. Растущее число МТС-устройств привело к исследованиям по борьбе с перегрузками и по уменьшению возможных коллизий, возникающих при установлении соединения по радиоканалу RACH [7–11].

В данной работе предлагается использование концепции виртуальных преамбул для более эффективного различения устройств LC-MTC и уменьшения коллизий.

2. Концепт виртуальной преамбулы

Процедура установления соединения предполагает на первом и втором шаге успешную отправку преамбулы и успешный прием ответного сообщения RAR (Random-Access-Response). Существующая схема установления соединения по радиоканалу RACH подразумевает наличие N_{pr} преамбул, N_{PRACH} физических

каналов PRACH (Physical Random Access Channel) для отправки преамбулы и канал PDCCH (Physical Downlink Control Channel) по нисходящей линии связи для отправки ответного сообщения RAR.

Новая схема установления соединения, предложенная для LC-MTC RACH в работе [2], использует вместо канала PDCCH канал EPDCCH (Enhanced Physical Downlink Control Channel), номер которого может быть дополнительным идентификатором в так называемой виртуальной преамбуле. Связка с номером канала EPDCCH увеличивает количество преамбул на коэффициент, равный количеству каналов EPDCCH [2]. В существующей схеме передачи преамбулы от MTC-устройства на базовую станцию для существующей процедуры установления соединения без привязки к индексу PRACH. При наличии одного идентификатора с номером преамбулы возможно в один момент времени уникально распознать не более N_{pr} преамбул. При получении двух преамбул с одним номером преамбулы базовая станция будет различать эти преамбулы по виртуальному идентификатору канала EPDCCH, и такой подход уменьшит возникающие коллизии, что послужит увеличению вероятности успешного установления соединения, уменьшению среднего времени установления соединения и других характеристик. В процессе установления соединения виртуальная преамбула на самом деле не передается, но неявно распознается между базовой станцией и устройствами LC-MTC. Поэтому устройства LC-MTC, использующие предложенную схему, могут быть совместимыми с теми MTC-устройствами, которые используют устаревшую схему установления соединения.

3. Вероятность успешного установления соединения

Исследована зависимость вероятности коллизии и вероятности успешного установления соединения от числа LC-MTC устройств в сети и количества EPDCCH каналов, по которым передаются RAR-сообщения. При введении дополнительного идентификатора в виде номера канала EPDCCH наблюдается уменьшение вероятности коллизии и увеличение вероятности успешного установления соединения для последующей передачи малых данных. Введение одного дополнительного канала EPDCCH уже влияет на характеристики. Для случая $N_{ch} = 2$ для 7 000 LC-MTC устройств вероятность коллизии преамбулы снижается в 1,7 раз. Рассмотрим введение двух дополнительных каналов EPDCCH. Для случая, когда в сети 30 000 LC-MTC устройств, вероятность коллизии падает со значения 0,95 до 0,62, при этом вероятность успешного установления соединения со значения 0,45 становится равной 0,99.

Подход с введением дополнительного идентификатора позволяет уменьшить вероятность коллизии, что влияет на все показатели, участвующие в процедуре установления соединения для MTC-устройств по радиоканалу RACH. Введение дополнительных идентификаторов увеличивает вероятность успешного установления соединения и уменьшает среднее время установления соединения.

Литература

1. Samouylov K. E., Gaidamaka Y. V., Gudkova I. A., Zaripova E. R., Shorgin S. Y. Baseline Analytical Model for Machine-Type Communications over 3GPP RACH in LTE-Advanced Networks / T. Czachórski et al. (Eds.): ISCIS 2016, CCIS 659. — Pp. 203–213.
2. Kim J. S., Lee S., Chung M. Y. Efficient Random-Access Scheme for Massive Connectivity in 3GPP Low-Cost Machine-Type Communications // IEEE Transactions on Vehicular Technology. — 2017. — Vol. 66, issue 7.
3. Islam M., Taha A. E., Akl S. A Survey of Access Management Techniques in Machine Type Communications, IEEE Communications Magazine. — May 2014. — Vol. 52, Issue 4. — Pp. 74–81.
4. 3GPP TS 22.368 v13.1.0, Technical Specification Group Services and System Aspects; Service Requirements for Machine-Type Communications (MTC); Stage 1 (Release 13), Dec. 2014.

5. Metnitzer A. Nokia Siemens Networks Machine2Machine Solution // In Proc. International Convention of MIPRO, May 2011. — Pp. 386–388.
6. Taleb T. and Kunz A. Machine Type Communications in 3GPP Networks: Potential, Challenges, and Solutions, IEEE Communications Magazine, Mar. 2012. — Vol. 50, Issue 3. — Pp. 178–184.
7. Hasan M., Hossain E., and Niyato D. Random Access for Machine-to-Machine Communication in LTE-Advanced Networks: Issues and Approaches // IEEE Communications Magazine. — June 2013. — Vol. 51, issue 6. — Pp. 86–93.
8. 3GPP TR 37.868 v11.0.0, Technical Specification Group Radio Access Network; Study on RAN Improvements for Machine-type Communications, Sep. 2011.
9. Lien S. Y., Chen K.-C., and Lin Y. Toward Ubiquitous Massive Accesses in 3GPP Machine-to-Machine Communications // IEEE Communications Magazine. — Apr. 2011. — Vol. 49, issue 4. — Pp. 66–74.
10. Ometov A., Sopin E., Gudkova I., Andreev S., Gaidamaka Y., Koucheryavy Y. Modeling Unreliable Operation of mmWave-Based Data Sessions in Mission-Critical PPDR Services // IEEE Access. — 2017. — Vol. 5, art. no. 8055435. — Pp. 20536–20544.
11. Зарипова Э. Р., Ардила Пинто А. Метод оценки времени установления соединения по радиоканалу случайного доступа // Вестник РУДН. Серия: Математика. Информатика. Физика. — 2017. — Т. 25, №. 1. — С. 9–18.
12. Ericsson R1-061369 TSG-RAN WG1 45 LTE random-access capacity and collision probability, Shanghai, China, May 8–12, 2006.

UDC 621.39

Success Access Probability Analysis Using Virtual Preambles Via Random Access Channel

I. E. Sinitsyn

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: iesinitsyn@gmail.com

Now rapid growth of number of the devices communicating among themselves in LTE (Long-Term Evolution) communication networks is observed, especially growth of number of machine-to-machine (M2M) devices should be noted because the number of devices will exceed 50 billion in recent years. To simplify low-cost network connection of a set of devices, communications of machine type (Machine-type communications, MTC) evolved in low-cost MTC (LC-MTC) in standards 3GPP. LC-MTC have to cope with intensive access through a set of narrow-band channels of random access (RACH) of LTE frequencies appropriated in the range effectively. With increase in quantity and density of placement of MTS of devices the scheme of random access for LC-MTC RACH has to be improved. In this paper we suggest to use the concept of virtual preambles in the procedure of random access and we carry out the success access probability analysis via a random access channel.

Key words and phrases: random access channel, collision, access success probability, low-cost MTC, preamble.

УДК 519.21

Модель надёжности однородной системы передачи данных облегчённого резервирования с произвольным распределением времени ремонта элементов

Г. Ж. К. Уанкпо, Д. В. Козырев

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклушто-Маклая, д. 6, Москва, Россия, 117198*

Email: gibsonhouankpo@yahoo.fr, kozyrev_dv@rudn.university

Рассматривается модель $\langle M_2/GI/1 \rangle$ системы облегчённого резервирования с экспоненциальной функцией распределения (ФР) времени безотказной работы (в.б.р.) и произвольной ФР времени ремонта её элементов. Получены явные аналитические выражения для стационарного распределения вероятностей состояний системы и для стационарной вероятности отказа системы. Проводится сравнение надёжности системы холодного резервирования и горячего резервированием через модель облегченного резервирования.

Ключевые слова: надёжность резервированных систем, стационарные вероятности, чувствительность, математическое моделирование.

1. Введение

Ранее [1–5] были проведены исследования надёжности систем передачи данных холодного резервирования. В текущей работе обобщаются результаты предыдущих исследований на случай облегченного резервирования системы $\langle M_2/GI/1 \rangle$ и проводится вычисление и сравнение стационарных характеристик надёжности для разных типов резервов. Целью работы является проведение аналитического моделирования дублированной системы $\langle M_2/GI/1 \rangle$ облегченного резервирования с произвольной ФР $B(x)$ и соответствующей плотностью распределения (ПР) $b(x)$ времени ремонта её элементов, и показателем распределения в.б.р. элементов с параметром $\lambda_i = \alpha + (1 - i)\gamma$; $i = 0, 1$.

2. Модель и аналитические результаты

Рассмотрим случайный процесс $v(t)$ — число отказавших элементов в момент времени t . Множество состояний процесса $E = \{0, 1, 2\}$.

Для описания поведения системы с помощью марковского процесса, введём дополнительную переменную $x(t) \in \mathbb{R}_+$ — время, затраченное в момент t , на ремонт отказавшего элемента. Получим двумерный процесс $(v(t), x(t))$ с расширенным пространством состояний $\epsilon = \{(0), (1, x), (2, x)\}$.

Обозначим $p_0(t)$ — вероятность того, что в момент времени t система находится в состоянии $i = 0$, $p_i(t, x)$ — плотность распределения (по непрерывной компоненте) вероятностей того, что в момент времени t система находится в состоянии i ($i = 1, 2$), и время затраченное на ремонт отказавшего элемента находится в интервале $(x, x + dx)$

$$p_0(t) = p\{v(t) = 0\},$$

$$p_1(t, x)dx = p\{v(t) = 1, x < x(t) < x + dx\},$$

$$p_2(t, x)dx = p\{v(t) = 2, x < x(t) < x + dx\}.$$

С помощью формулы полной вероятности перейдем к выводу системы уравнений Колмогорова, а с предельным переходом при $\Delta \rightarrow 0$ и в предположении что процесс имеет стационарное распределение при $t \rightarrow \infty$, получаем систему уравнений баланса в виде:

$$\begin{cases} \lambda_0 \cdot p_0 = \int_0^\infty p_1(x) \cdot \delta(x) dx \\ \frac{\partial p_1(x)}{\partial x} = -(\lambda_1 + \delta(x)) \cdot p_1(x) \\ \frac{\partial p_2(x)}{\partial x} = -\delta(x) \cdot p_2(x) + \lambda_1 \cdot p_1(x), \end{cases}$$

где $\delta(x) = \frac{b(x)}{1-B(x)}$ — условная ПР остаточной длительности ремонта элемента, находящегося в ремонте время t (интенсивность восстановления). Граничное условие:

$$p_1(0) = \lambda_0 \cdot p_0(x) + \int_0^\infty p_2(x) \cdot \delta(x) dx$$

Решая полученную СДУК, используя метод вариации постоянной, получаем стационарные вероятности состояний системы облегченного резервирования:

$$p_0 = C_1 \cdot (\lambda_0^{-1} \cdot \tilde{b}(\lambda_1)), \quad p_1 = C_1 \cdot \left(\frac{1 - \tilde{b}(\lambda_1)}{\lambda_1} \right), \quad p_2 = C_1 \cdot \left(\frac{\rho^{-1} - (1 - \tilde{b}(\lambda_1))}{\lambda_1} \right).$$

где $\rho^{-1} = \frac{EB}{EA} = b \cdot \lambda_1$. и $C_1 = \frac{\lambda_0 \cdot \lambda_1}{\lambda_0 \cdot \rho^{-1} + \lambda_1 \cdot \tilde{b}(\lambda_1)}$.

Очевидно, имеется зависимость стационарных вероятностей состояний системы от вида распределений времени ремонта, однако анализ и численные исследования показывают, что при «быстром» восстановлении она становится исчезающе малой. Также видно, что для всех распределений, вероятность отказа системы p_2 при горячем резервировании будет больше чем при холодном, то есть самой надежной является модель с холодным типом резерва. При $\gamma = 0$, система передачи данных находится в холодном резерве, а при $\gamma = \alpha$ — в горячем резерве. В табл. 1 приведены значения стационарных вероятностей отказа системы для разных ФР.

Таблица 1
Стационарные вероятности отказа p_2 системы $\langle M_2/GI/1 \rangle$, рассчитанные аналитически при разных значениях модельного параметра $\rho = 1, 10, 100, 1000$

ρ	GI	$EXP\left(\beta = \frac{1}{EB}\right)$		$WB\left(k = \frac{1}{2}, \beta = \frac{k}{EB}\right)$		$PAR\left(k = 2, \frac{EB(k-1)}{k}\right)$	
		$\gamma = 0$	$\gamma = \alpha$	$\gamma = 0$	$\gamma = \alpha$	$\gamma = 0$	$\gamma = \alpha$
1	P_2	0.3333333	0.4	0.3960184	0.4937942	0.3070999	0.3628093
10	P_2	0.009009008	0.01639344	0.02036197	0.03709043	0.009744243	0.01773251
100	P_2	$9.9 \cdot 10^{-5}$	0.00019604	0.000285907	0.0005661541	0.0001580345	0.0003129401
1000	P_2	$9.99 \cdot 10^{-7}$	$1.996 \cdot 10^{-6}$	$2.983 \cdot 10^{-6}$	$5.96 \cdot 10^{-6}$	$4.31 \cdot 10^{-6}$	$8.61 \cdot 10^{-6}$

3. Заключение

В работе были получены явные аналитические выражения для стационарного распределения вероятностей состояний системы облегченного дублирования и для стационарной вероятности отказа системы. Полученные формулы показывают наличие явной зависимости этих характеристик от вида ФР времени ремонта её элементов. Модель холодного резервирования является самой надежной для всех функций распределений.

Литература

1. Уанкпо Г.Ж.К., Козырев Д.В. Анализ чувствительности характеристик надёжности модели резервирования системы передачи данных к виду распределений времени безотказной работы и ремонта её элементов // Информационно-телекоммуникационные технологии и математическое моделирование высоко-технологичных систем (ИГТММ-2017). Материалы Всероссийской конференции с международным участием. – М.: РУДН, 2017. — С. 55–58.
2. Houankpo H. G. K., Kozыrev D. V. Sensitivity Analysis of Steady State Reliability Characteristics of a Repairable Cold Standby Data Transmission System to the Shapes of Lifetime and Repair Time Distributions of its Elements // In: K. E. Samouilov, L. A. Sevastianov, D. S. Kulyabov (eds.): Selected Papers of the VII Conference “Information and Telecommunication Technologies and Mathematical Modeling of High-Tech Systems”, Moscow, Russia, 24-Apr-2017, published at <http://ceur-ws.org/Vol-1995/2>.
3. Kozыrev D., Kimenchezhi V., Houankpo H. G. K. Reliability Calculation of a Redundant Heterogeneous System with General Repair Time Distribution // Прикладные проблемы в теории вероятностей и математической статистике в области телекоммуникаций = Applied problems in theory of probabilities and mathematical statistics into telecommunications. Труды XI Международного семинара / под ред. Д. Аранити, К. Е. Самуйлова, С. Я. Шоргина. — М: РУДН, 2017. — С. 12.
4. Уанкпо Г.Ж.К., Козырев Д.В. Исследование чувствительности характеристик надёжности резервированной системы передачи данных к виду распределений времени между отказами и восстановлениями элементов системы // Молодежная научная школа по прикладной теории вероятностей и телекоммуникационным технологиям (АРТСТ–2017) = 2nd International School on Applied Probability Theory and Communications Technologies (АРТСТ–2017): материалы молодежной научной школы. Россия, Москва, 23–27 октября 2017 г.; под общ. ред. К. Е. Самуйлова, Е. А. Кучерявого, А. Н. Дудина. — Москва: РУДН, 2017. — С. 299–303.
5. Уанкпо Г. Ж. К., Козырев Д. В. Анализ чувствительности характеристик надёжности модели дублированной системы передачи данных к виду распределений времени безотказной работы и ремонта её элементов / Материалы 19-й Международной научной конференции «Распределенные компьютерные и телекоммуникационные сети: управление, вычисление, связь» (DCCN-2016, Москва). — М.: РУДН, 2016. Том 3 / под общей ред. В.М. Вишневого и К. Е. Самуйлова. — 2016. — С. 473–480.

UDC 519.21

Reliability Model of a Homogeneous Warm-Standby Data Transmission System with General Repair Time Distribution

H. G. K. Houankpo, D. V. Kozыrev

Department of Applied Probability and Informatics

Peoples' Friendship University of Russia

6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation

Email: gibsonhouankpo@yahoo.fr, kozыrev_dv@rudn.university

We consider a mathematical model of a repairable warm-standby data transmission system and study its reliability. Explicit analytical expressions are obtained for the steady state system probabilities and for the stationary probability of the system-level failure. Comparison of the system-level reliability under cold and hot standby was conducted.

Key words and phrases: system-level reliability, steady state probabilities, warm standby, mathematical modeling, sensitivity.

УДК 519.872

Анализ временных показателей бизнес-процессов

В. А. Халина

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: khalina_va@rudn.university

В работе рассматривается моделирование эталонных сквозных бизнес-процессов в виде открытой ВСМР-сети. Для этого использовались методы аппарата экспоненциальных сетей массового обслуживания с неоднородными заявками. Кроме того, приводится метод численного анализа, в частности, формула среднего времени нахождения заявки в узле сети, с помощью которой вычисляются стандартные бизнес-показатели TM Forum.

Ключевые слова: теория массового обслуживания, моделирование бизнес-процессов, анализ показателей эффективности.

1. Введение

В современных условиях конкуренция на рынке услуг постоянно возрастает. Появляются новые бизнес-модели, технологии усложняются, расширяется спектр услуг. Сфера телекоммуникационных услуг также не стоит на месте. Поэтому, чтобы реализовать успешный бизнес в данной области, необходимо сокращать расходы и как можно эффективнее осуществлять цифровое преобразование. В этом помогает тщательный анализ бизнес-процессов.

Так, большое количество клиентских запросов и обслуживание в режиме реального времени позволяют принять бизнес-процессы телекоммуникационной компании как хороший предмет для анализа с помощью теории массового обслуживания. Более того, этому анализу способствует обширный набор стандартов и рекомендаций, разработанных международной ассоциацией в области телекоммуникаций TM Forum [1].

Поэтому, актуальна задача построения математической модели бизнес-процессов и приведение численного анализа ее временных характеристик объединенной модели бизнес-процессов.

2. Основная часть

В реалиях бизнеса ни один из процессов не протекает автономно от других. Их потоки пересекаются на том или ином уровне. Поэтому, основываясь на методе, описанном в [2], построена модель, показанная на (рис. 1) и объединяющая 5 эталонных сквозных бизнес-процессов карты eTOM [3]. Модель построена в виде экспоненциальной сети с неоднородными заявками, т.н. ВСМР-сети [4]. Рассмотрены процессы «Запрос-Решение», «Жалоба-Решение», «Сбой-Решение», «Прекращение-Подтверждение», «Запрос-Изменение». Таким образом, модель наиболее адекватно отражает процессы телекоммуникационной компании и дает возможность оценить уровень производительности процессов.

Опираясь на [2], разработан метод анализа временных показателей бизнес-процессов, который основывается на формуле (1):

$$\bar{T}_i = \begin{cases} \frac{c_i \mu_i}{(c_i \mu_i - \lambda_i)^2} \frac{\rho_i^{c_i}}{c_i!} \left(\sum_{m=0}^{c_i-1} \frac{\rho_i^m}{m!} + \frac{\rho_i^{c_i}}{c_i!} \frac{1}{1 - \frac{\rho_i}{c_i}} \right) + \frac{1}{\mu_i}, & i \in M_{FCFS}, \\ \frac{1}{\mu_i}, & i \in M_{IS}, \end{cases} \quad (1)$$

где c_i — число приборов на i узле, μ_i — интенсивность обслуживания на i узле, λ_i — интенсивность входящего потока на i узел, ρ_i — загрузка i узла, m — число узлов.

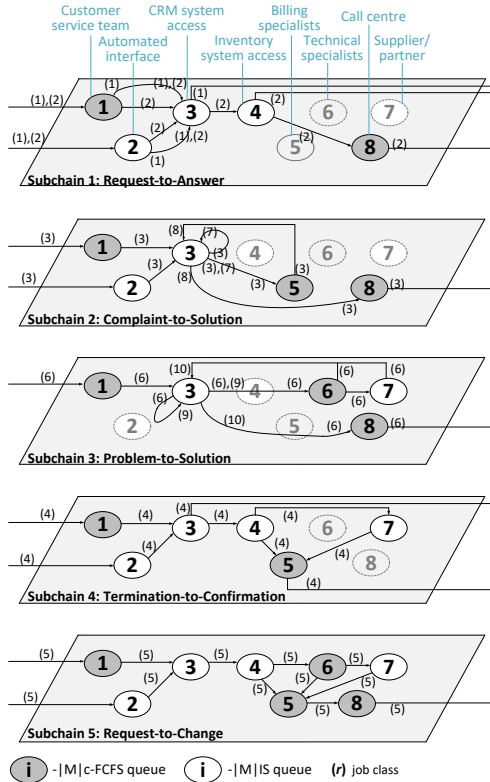


Рис. 1. Объединенная модель сквозных бизнес-процессов в виде сети ВСМР

Формула (1) позволяет получить выражения для вычисления значений стандартных бизнес-показателей TM Forum [5].

3. Заключение

Таким образом, при помощи объединения нескольких бизнес-процессов в одной сети, учтено совместное использование ресурсами, происходящее на функциональных уровнях компании, например, в биллинговой группе или у операторов call-центра.

Разработан подход к оценке некоторых важных бизнес-показателей TM Forum с использованием сети ВСМР, которая объединяет модели нескольких стандартных сквозных бизнес-потоков eTOM. Этот метод может применяться для анализа набора бизнес-процессов конкретного поставщика услуг.

В ходе дальнейших исследований поставлена задача проверки результатов анализа путем сравнения с результатами имитационного моделирования.

Литература

1. TM Forum Framework, <https://www.tmforum.org/tm-forum-framework>
2. Samouylov K., Gaidamaka Y., Zaripova E. Analysis of Business Process Execution Time with Queueing Theory Models // Information Technologies and Mathematical Modelling - Queueing Theory and Applications / Dudin A., Gortsev A., Nazarov A., Yakupov R. (eds.); Communications in Computer and Information Science. — 2016. — Vol. 638. — Pp. 315–326. doi:10.1007/978-3-319-44615-8_28.
3. TM Forum: Business process framework (eTOM) End-to-end business flows. Framework Standard GB921 Addendum E. R16.5.1, Vol. 16.5.2, 2017.
4. Baskett F., Chandy K. M., Muntz R. R., Palacios F. G. Open, Closed, and Mixed Networks of Queues with Different Classes of Customers // Journal of the ACM. — 1975. Vol. 22, Pp. 248–260. doi:10.1145/321879.321887.
5. TM Forum: TM Forum Metrics Definitions. Framework Best Practice GB988. R16.5.0, Vol. 4.0.1, 2016.

UDC 519.872

Business Processes' Time Metrics Analysis

V. A. Khalina

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: khalina_va@rudn.university

The paper refers to the end-to-end business processes modeling using a BCMP network. In this case, the methods of exponential queueing networks with heterogeneous jobs were used. In addition, the method of numerical analysis is given, in particular, the formulae for the job average hold time in a network node, by which the TM Forum's standard business indicators are calculated.

Key words and phrases: queueing theory, business process modelling, metrics analysis.

УДК 519.872, 519.217

К анализу системы массового обслуживания с ресурсами, функционирующей в случайном окружении

В. В. Цурлуков*, И. С. Зарядов*[†], Т. А. Милованова*

* Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198

[†] Институт проблем информатики
Федеральный исследовательский центр «Информатика и управление» РАН
ул. Вавилова, д. 44, к. 2, Москва, Россия, 119333

Email: dober.vvt@gmail.com, zaryadov_is@rudn.university, milovanova_ta@rudn.university

Строится математическая модель системы, предоставляющей поступающим заявкам доступ к ресурсам и функционирующей в случайном окружении.

Ключевые слова: система массового обслуживания, случайное окружение, ММРР, предоставление ресурсов.

1. Введение

Математическая модель анализа функционирования современных телекоммуникационных систем должна учитывать воздействие внешних факторов, что реализуется в рамках теории массового обслуживания (теории телетрафика) [1, 2] с помощью процессов поступления и/или обслуживания, управляемых внешним случайным процессом. Применение процессов поступления/обслуживания, управляемых внешним марковским процессом (Markov modulated arrival process (ММАР), Markov modulated service process (ММСП)) [3–8], позволяет в рамках исследования построить не только адекватную математическую модель, но и получить хорошие аналитические выражения.

В настоящее время актуальными являются задачи, в рамках решения которых моделируются и анализируются системы, когда поступающим запросам помимо обслуживания требуется еще и предоставление некоторых дополнительных ресурсов фиксированного или переменного объема [9–12].

В данной работе представлено описание системы, в которой процесс поступления требований на ресурсы фиксированного объема управляется внешним марковским процессом. Кроме того, интенсивность обслуживания и общий объем ресурсов системы также регулируются внешним марковским процессом (отличным от марковского процесса, регулирующего поступление запросов в систему). Начальные этапы данного исследования были представлены в [13].

2. Основная часть

Рассматриваемую систему, в которой поступление и обслуживание управляются внешними марковскими процессами, согласно нотации Башарина–Кендалла [1, 2] можно обозначить следующим образом — $MMPP_2|MMSP_2[n|r][R_1, R_2]$. Система состоит из $1 \leq n < \infty$ однородных приборов, время обслуживания на каждом из которых подчинено экспоненциальному закону, и накопителя емкости $r \leq \infty$.

Случайное окружение представлено двумя марковскими процессами (МП) $\eta_1(t)$ и $\eta_2(t)$. Здесь $\eta_1(t)$ — это МП, управляющий поступлением требований и пребывающий в одном из двух возможных состояний. Если $\eta_1(t)$ находится в состоянии 1, то требования поступают согласно пуассоновскому закону с интенсивностью λ_1 и каждому требованию нужен фиксированный объем ресурсов размера k_1 . В состоянии 2 МП $\eta_1(t)$ требования, которым нужен тоже фиксированный объем ресурсов, но уже размера k_2 , поступают с интенсивностью λ_2 .

Марковский процесс, который управляет объемом ресурсов, предоставляемых системой, а также интенсивностью обслуживания требований на приборах, тоже может пребывать только в одном из двух состояний. Если МП $\eta_2(t)$ находится в состоянии 1, то максимальный объем ресурсов системы равен $R_1 < \infty$, время обслуживания требования на любом приборе подчинено экспоненциальному распределению с интенсивностью μ_1 . Если же состояние МП $\eta_2(t) = 2$, то требования на приборах обслуживаются с интенсивностью μ_2 и максимальный объем ресурсов системы равен $R_2 = \infty$. Переходы управляющих МП $\eta_1(t)$ и $\eta_2(t)$ из одного состояния в другое определяются соответствующими инфинитезимальными матрицами $\Lambda = (\lambda_{ij})_{i,j=1,2}$ и $M = (\mu_{ij})_{i,j=1,2}$.

Случайный процесс $\zeta(t) = \{\xi_1(t), \xi_2(t), R(t), \eta_1(t), \eta_2(t)\}$ описывает функционирование системы. Здесь $\xi_1(t) = (\xi_{1s}(t), \xi_{1q}(t))$ — число требований на ресурсы размера k_1 , находящихся на приборах ($\xi_{1s}(t)$) и в накопителе ($\xi_{1q}(t)$) в момент времени t . Соответственно, $\xi_2(t) = (\xi_{2s}(t), \xi_{2q}(t))$ — число требований на ресурсы размера k_2 на приборах ($\xi_{2s}(t)$) и в накопителе ($\xi_{2q}(t)$) в момент времени t . Через $R(t)$ обозначается доступный объем ресурсов всей системы в момент времени t . Если МП η_2 находится в состоянии 1, то $R(t) = \max(0, R_1 - k_1\xi_1(t)\mathbf{1} - k_2\xi_2(t)\mathbf{1})$. Если МП η_2 находится в состоянии 2, то $R(t) = R_2 = \infty$.

В рамках решаемой задачи будем считать, что максимального конечного объема ресурсов R_1 достаточно для обслуживания всех требований на имеющихся приборах (невозможна ситуация, когда простаивает прибор из-за нехватки ресурсов), то есть $R_1 \geq n \cdot \max(k_1, k_2)$.

Также будем считать, что, если текущий объем ресурсов $R(t)$ меньше, чем требуется поступающей заявке, то данная заявка теряется. Также возможны потери заявок, уже находящихся в системе, из-за перехода системы из состояния с неограниченным объемом ресурсов R_2 в состояние с ограниченным объемом $R(t) = R_1 - k_1\xi_1(t)\mathbf{1} - k_2\xi_2(t)\mathbf{1}$.

3. Заключение

В работе построена математическая модель системы, функционирующей в случайной среде, с выделением ресурсов поступающим заявкам, сформулированы основные задачи исследования, представлены полученные результаты

Благодарности

Публикация подготовлена при поддержке Программы РУДН «5-100» и при финансовой поддержке РФФИ в рамках научных проектов № 18-07-00692, 16-07-00766.

Литература

1. Бочаров П. П., Печинкин А. В. Теория массового обслуживания, М.: РУДН, 1995.
2. Башарин Г. П. Лекции по математической теории телетрафика, М.: Изд. РУДН, 3-е изд., испр. и доп., 2009.
3. Neuts M. F. A versatile Markovian point process // Journal of Applied Probability. — 1979. — Vol. 16, no. 4. — Pp. 764–779.
4. Fisher W., Meier-Hellstern K. S. The Markov-modulated Poisson process (MMPP) cookbook // Performance Evaluation. — 1993. — Vol. 18, no. 2. — Pp. 149–171.
5. Ng Chee-Hoc, Soong Boon-Hee. Queueing Modelling Fundamentals with Applications in Communication Networks. 2nd Edition, John Wiley & Sons Ltd, 2008.

6. Цурлуков В. В., Крупко О. С., Зарядов И. С. Входящий поток, управляемый целью Маркова (ММАР). Использование в задачах теории телеграфика // Материалы Всероссийской конференции с международным участием «Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем». Москва, РУДН, 24–28 апреля 2017, М.: РУДН. — С. 66–68.
7. Zaryadov I., Korolkova A., Kulyabov A., Milovanova T., Tsurulukov V. The Survey on Markov-Modulated Arrival Processes and Their Application to the Analysis of Active Queue Management Algorithms, DCCN 2017 // Communications in Computer and Information Science. — 2017. — Vol. 700. — Pp. 417–430.
8. Abaev P., Razumchik R. Stationary distribution of $MMPP|D|1|R$ queue with bi-level hysteric policy // XXXII International Seminar on Stability Problems for Stochastic Models. June 16–21, 2014, Norwegian University of Science and Technology, Trondheim, Norway, 2014. — Pp. 113–114.
9. Naumov V., Samouylov K., Yarkina N., Sopin E., Andreev S., Samuylov A. LTE Performance Analysis Using Queuing Systems with Finite Resources and Random Requirements // 7th International Congress on Ultra Modern Telecommunications and Control Systems ICUMT-2015, October 6–8, 2015, Brno, Czech Republic, 2015. — Pp. 100–103.
10. Sopin E., Vikhrova O., Samouylov K. LTE network model with signals and random resource requirements // 9th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops, ICUMT-2017, IEEE, 2017. — Pp. 101–106. <http://ieeexplore.ieee.org/document/8255155/>
11. Самуйлов К. Е., Сопин Э. С., Шоргин С. Я. Система массового обслуживания с ограниченными ресурсами и сигналами для анализа показателей эффективности беспроводных сетей // Информатика и ее применения. — 2017. — Т. 11, № 3. — С. 99–105.
12. Сопин Э. С., Самуйлов К. Е. Рекуррентный алгоритм вычисления вероятностных характеристик для СМО с ограниченными ресурсами и случайными требованиями // Материалы Всероссийской конференции с международным участием «Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем». Москва, РУДН, 24–28 апреля 2017. — М.: РУДН, 2017. — С. 277–280.
13. Цурлуков В. В. Математическая модель системы с ресурсами в случайном окружении, Материалы международной молодежной научной школы «Молодежная научная школа по прикладной теории вероятностей и телекоммуникационным технологиям (АРТСТ-2017) — 2nd International School on Applied Probability Theory and Communications Technologies». Россия, Москва, 23–27 октября 2017 г. — М.: РУДН, 2017. — С. 318–320.

UDC 519.872, 519.217

To the Analysis of the Queuing System Operating in the Random Environment with Resource Allocation

V. V. Tsurulukov*, I. S. Zaryadov*[†], T. A. Milovanova*

* *Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

[†] *Institute of Informatics Problems, FRC CSC RAS
44-2 Vavilova str., Moscow 119333, Russian Federation*

Email: dober.vvt@gmail.com, zaryadov_is@rudn.university, milovanova_ta@rudn.university

The mathematical model of the system which operates in the random environment and provides the incoming tasks with access to some value of resources is introduced.

Key words and phrases: queuing system, random environment, Markov modulated Poisson process, Markov modulated service process, resource allocation.

**Сети связи следующего поколения: управление,
качество, архитектура**

**Next-generation networks: management, quality,
architecture**

УДК 004.4

Модель для анализа показателей эффективности ядра сети 5G с виртуализацией и гистерезисным управлением

П. О. Абаев*[†], А. С. Царев*

* *Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

[†] *Федеральный исследовательский центр «Информатика и управление» РАН
ул. Вавилова, д. 44, к. 2, Москва, Россия, 119333*

Email: abaeв_po@rudn.university, atsarev@sci.pfu.edu.ru

В настоящее время идет активная работа по стандартизации, исследованиям и разработке сетей мобильной связи пятого поколения (5G). Программно-конфигурируемые сети (SDN) и виртуализация сетевых функций (NFV) — ключевые технологии, которые часто упоминаются в исследованиях. Архитектура, сочетающая в себе оба этих подхода, увеличивает гибкость, производительность и эффективность сети. Разделение уровня данных и уровня управления в SDN позволяет осуществлять гибкое развертывание новых узлов сети. В данной работе строится модель виртуализированного ядра сети и предлагается алгоритм управления процессом разворота/сворота виртуальных ресурсов.

Ключевые слова: виртуализация, NVF, SDN, ПКС, 5G, СМО, ТМО, гистерезисное управление.

1. Введение

Деятельность по разработке и стандартизации сетей 5G все еще находится на ранней стадии, но тем не менее уже признано, что системы, основанные на технологии 5G будут широко использоваться гетерогенные сетевые технологии, опирающиеся на парадигмы виртуализации сетевых функций (NFV) и программно-конфигурируемых сетей (SDN) [1]. Использование технологий NFV и SDN позволяет создавать новые сервисы с существенно более высокими скоростями передачи данных и меньшими задержками [2]. SDN и NFV были выбраны в качестве ключевых факторов эволюции сети, направившие процесс стандартизации 5G в новом направлении [3].

С использованием NFV [4] станет возможна виртуализация физических ресурсов сети, позволив таким образом масштабировать логические ресурсы сетевой инфраструктуры. В частности, предполагается, что ядро сети будет полностью виртуализированно. Ядро сети является одной из наиболее важных частей сети. Главная цель при использовании NFV заключается в виртуализации функций ядра сети. Основной компонентой ядра сети является EPC. Виртуализированный EPC (vEPC) можно представить в виде модели, где устаревшее сетевое оборудование считается зарезервированным блоком серверов, а экземпляры VNF включаются и выключаются в соответствии с количеством представленных запросов на работу [5].

2. Основная часть

Построим математическую модель рассмотренного виртуализированного ядра сети 5G. На многолинейную систему с C обслуживающими приборами и накопителем конечной емкости R поступает пуассоновский поток заявок с интенсивностью $\lambda, 0 \leq \lambda < \infty$. Длительность обслуживания заявки распределена экспоненциально с параметром $\mu, 0 \leq \mu < \infty$. Каждой заявке для обслуживания требуется один прибор. Если в накопителе нет места, то пришедшая заявка теряется.

Для увеличения эффективности построенной модели предлагается использование гистерезисного механизма управления. В системе имеются два порога: L и H . Система может находиться в одном из четырех состояний:

- простой дополнительных приборов (*idle*);
- разворот дополнительных приборов (*up*);
- сворачивание дополнительных приборов (*down*);
- все приборы развернуты (*full*).

Система начинает свою работу в состоянии *idle*, обработкой заявок занимаются c_0 приборов (традиционная часть ядра), а $k = C - c_0$ дополнительных приборов простаивают. Если число заявок в накопителе превысит порог H , то простаивающие приборы начинают обработку заявок через интервал времени, который распределен экспоненциально с параметром $\theta, 0 \leq \theta < \infty$, и система перейдет в состояние *up*. Если все приборы в системе будут развернуты – система перейдет в состояние *full*. Как только число заявок в накопителе станет меньше порога L , окончившие обслуживание заявок приборы прекращают свою работу, пока число работающих приборов не уменьшится до c_0 , а система переходит в состояние *down*. На рис. 1 изображена описанная математическая модель.

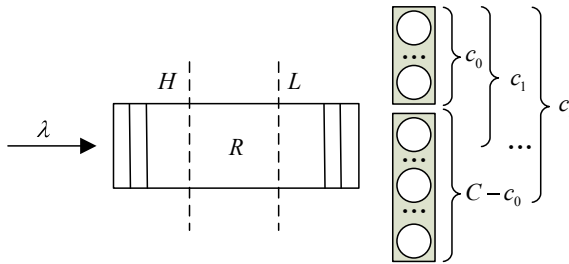


Рис. 1. Схема системы массового обслуживания

Составной случайный процесс $X(t) = (s(t), n(t), m(t))$, где $s = \{idle, up, down, full\} = \{i, d, u, f\}$ – состояние системы, $0 \leq n \leq k$ – количество активных приборов, $0 \leq m \leq R + C = K$ – количество заявок в системе. Этот процесс по построению является марковским и описывает поведение системы над пространством состояний:

$$\chi(t) = \left\{ (s, n, m) : \left(\begin{array}{l} s = i, n = 0, \quad 0 \leq m \leq H, \\ s = d, \quad 0 \leq n < k, \quad c_n \leq m \leq H, \\ s = u, \quad 0 \leq n < k, \quad L < m < R, \\ s = f, \quad n = k \end{array} \right) \right\}.$$

3. Заключение

В работе построена математическая модель виртуализированного ядра сети 5G с гистерезисным управлением. Разработана имитационная модель описанной модели и проанализированы преимущества системы с управлением перед системой без управления.

Благодарности

Публикация подготовлена при поддержке Программы РУДН «5-100».

Литература

1. Andrews J. G., Buzzi S., Choi W., Hanly S. V., Lozano A., Soong A., Zhang J. What will 5G be? // *IEEE Journal on Selected Areas in Communications*. — June 2014. — Vol. 32, no. 6. — Pp. 1065–1082.
2. Agyapong P. K., Iwamura M., Staehle D., Kiess W., Benjebbour A. Design considerations for a 5G network architecture // *IEEE Communications Magazine*. — Nov 2014. — Vol. 52, no. 11. — Pp. 65–75.
3. Ordonez-Lucena J., Ameigeiras P., Lopez D., Ramos-Munoz J., Lorca J., Figueira J. Network slicing for 5G with SDN/NFV: Concepts, architectures, and challenges // *IEEE Communications Magazine*. — May 2017. — Vol. 55, no. 5. — Pp. 80–87.
4. ETSI, Network Functions Virtualization (NFV): Use Cases // *ETSI GS NFV 001, v.1.1.1* – Oct. 2013.
5. Yi Ren, Tuan Phung-Duc, Zheng-Wei Yu, Yuh-Cheng Chen. Design and Analysis of Dynamic Auto Scaling Algorithm (DASA) for 5G Mobile Networks // *IEEE GLOBECOM, Washington, DC USA* – Dec. 2016. — Pp. 1–17.

UDC 004.4

Design and Performance Evaluation of Virtualized 5G Network Core Model with Implemented Hysteresis Control

P. O. Abaev*[†], A. S. Tsarev*

** Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

*[†] Institute of Informatics Problems, FRC CSC RAS
44-2 Vavilova Str., Moscow 119333, Russian Federation*

Email: abaev_po@rudn.university, atsarev@sci.pfu.edu.ru

The process of active work on standardization, research and development of 5G telecommunication networks is still ongoing. Software-Defined Networks (SDN) and Network Functions Virtualization (NFV) approaches are key technologies that are often referred to in 5G white papers. The flexibility, productivity and efficiency of the networks increased combining both approaches. Separating the data plane and the control plane in SDN allows to deploy new nodes in more flexible way. In this paper a model of the virtualized 5G network core is designed and an deployment control algorithm is proposed.

Key words and phrases: NFV, VNF, 5G, SDN, Hysteresis Control, Queuing System, Performance Evaluation.

УДК 519.254.1

Прямое взаимодействие пользователей в мобильных социальных сетях

С. Д. Андреев

*Федеральный исследовательский центр «Информатика и управление» РАН
ул. Вавилова, д. 44, к. 2, Москва, Россия, 119333*

Email: serge.andreev@gmail.com

Рассмотрены предпосылки и последствия развития мобильных социальных сетей, а также использования в них прямого взаимодействия между пользователями.

Ключевые слова: мобильные социальные сети, прямое взаимодействие устройств.

1. Введение

Использование персональных мобильных устройств стало неразрывно связано с нашей повседневной жизнедеятельностью благодаря повсеместной распространенности доступных средств беспроводного доступа. В результате, мобильная связь быстро становится одной из первоочередных глобальных потребностей, которая подкрепляется тем фактом, что еще в 2015 году смартфонами владело уже 80% Интернет-пользователей. На сегодняшний день, каждый абонент в среднем выполняет на своем мобильном устройстве более 1500 задач в неделю, начиная с раннего утра при проверке персональных электронных писем и профиля Facebook. Современные статистические данные свидетельствуют об использовании смартфона своим владельцем более трех часов в день. В дополнение к смартфонам, которые позволяют производить, обмениваться и отправлять различный цифровой контент, также активно применяется другое персональное оборудование, такое как портативные компьютеры, планшеты и носимые устройства [1].

Не удивительно, что возникающий как следствие рост объемов мобильного трафика уже превышает 70% и прогнозируется к увеличению еще в 10 раз за ближайшие 5 лет, что составляет существенную нагрузку на современные системы мобильной связи. В частности, значительная доля передаваемых данных связана с созданием абонентами собственного мультимедиа-контента, который далее поступает в социальные сети [2]. Развитие сетей связи 5G предполагает одной из своих целей обслуживание такой лавинной загрузки, что достигается применением более высоких частот, «массивных» многоантенных систем и сверхплотного размещения сетевой инфраструктуры, приводя к существенному повышению доступной емкости сети [3]. С другой стороны, развитие технологий производства мобильных устройств снабжает их все более совершенными средствами связи и обработки информации, приводя к появлению интеллектуального абонентского оборудования, которое участвует в управлении сетевым взаимодействием наравне с сетевой инфраструктурой.

Интеллектуальные пользовательские устройства, снабженные различными продвинутыми технологиями радиосвязи, а также более надежными механизмами хранения данных и борьбы с интерференцией, играют все более заметную роль в управлении быстро формирующимися системами 5G. В частности, прямое взаимодействие «устройство-устройство» (device-to-device, D2D) позволяет повысить эффективность обмена пользовательской информацией. В результате, возникает принципиальный переход от ставшего уже аксиоматическим подхода к управлению сетевой инфраструктурой со стороны инфраструктуры мобильного оператора к более перспективным способам, предполагающим вовлечение абонентских терминалов. В итоге, контроль за сетевым взаимодействием смещается от ядра сети к ее периферии, что, в свою очередь, меняет концепцию использования современных телекоммуникационных технологий. Действительно, тогда как в системах предыдущих поколений

для передачи голосовых потоков весь трафик проходил исключительно через сетевую инфраструктуру, в современных системах, ориентированных на отправку данных, расположенные близко друг к другу пользователи имеют возможность обмениваться контентом напрямую.

2. Предпосылки возникновения и развития мобильных социальных сетей

Прямое взаимодействие между абонентскими терминалами принципиально отличается от централизованного функционирования тем, что оно неразрывно связано с социальным поведением пользователей и вызвано их непосредственными нуждами. Интернет-форумы и чат-группы стали первыми в истории примерами социальных сетей, возникновение которых было обусловлено потребностью в удаленном общении. Позднее к ним присоединились многочисленные файлообменные сети, поддерживаемые соответствующими социальными сервисами [4]. Сравнительно недавно, в связи с увеличением технологических возможностей абонентских терминалов, их владельцы стали все больше переходить на пользование мобильными услугами. Это стало возможным благодаря более точному позиционированию (при помощи сотовых технологий и системы GPS), наличию контекстной информации (получаемой через сенсоры), средствам связи ближнего радиуса действия, а также повсеместному доступу в сеть Интернет.

На сегодняшний день, миллиарды людей получают в реальном масштабе времени свободный доступ к широкому множеству социальных сетей, таких как Facebook, LinkedIn, Twitter, Instagram и Google+, при помощи всевозможных персональных устройств. Из социологии известно, что социальные контакты людей естественным образом кластеризуются, поскольку жизнь человека подчинена определенной каждодневной рутине: взаимодействия между людьми и их социальные взаимоотношения в высокой степени локальны, так как организованы вокруг некоторого общего физического пространства. Однако с ростом возможностей для удаленного общения, мобильные социальные сети смещают «центр тяжести» таких контактов из физических в виртуальные сообщества (т.е., группы людей, обменивающихся информацией удаленно). Эти сообщества образуются, в основном, индивидами, сгруппированными в соответствии с определенной социальной структурой, имеющей специфические (устойчивые) механизмы взаимодействия.

Тогда как удаленное взаимодействие посредством централизованных сетей передачи данных действительно оказало значительное влияние на продуктивность и информированность людей, оно также наложило ряд ограничений на естественность человеческого общения. Основным сдерживающим фактором в развитии виртуальных сообществ (и для полезности удаленных социальных контактов) является тот факт, что пользователи не осведомлены о богатых социальных возможностях, имеющихся в их непосредственном окружении (например, предпосылки для новых знакомств, как деловых, так и развлекательных, коллективных игр, обмена контекстными мультимедиа-данными и др.). В подтверждение этого обстоятельства недавние исследовательские результаты свидетельствуют о том, что удаленное взаимодействие действительно стало препятствием по отношению к непосредственным социальным контактам, которые, вообще говоря, являются наиболее естественной и исторически обусловленной формой деятельности человека.

Таким образом, перспективные технологии связи 5G должны предоставить средства для эффективного сопряжения физического и виртуального сообществ [5], позволяя своим пользователям взаимодействовать с «подходящими» партнерами (т.е., с такими, которые имеют общие с ними интересы), а также с объектами, находящимися в непосредственной близости. Это связано с предоставлением возможностей для виртуального участия в совместных социальных контактах, учитывающих личные предпочтения. Такое направление развития не только создает предпосылки для более глубокого вовлечения людей в процесс общения, но также позволяет усилить и само социальное взаимодействие. Последнее становится реальным благодаря созданию устойчивых связей между людьми и их подходящими навыками,

позволяя упрочнить чувство сплоченности в условиях городской обстановки, где ощущение одиночества и изолированности уже давно стало значительным поводом для беспокойства.

3. Выводы

В свете вышеизложенного, следует отметить неоспоримые преимущества применения мобильных социальных сетей с возможностью прямого взаимодействия их пользователей. При этом, также возникают новые трудности, связанные с автономным обнаружением подходящих социальных контактов, а также с идентификацией абонентов в реальном масштабе времени и доверенным управлением контентом в общем окружении, включая обеспечение качества обслуживания и восприятия услуг. Таким образом, возникает задача унификации этой перспективной формы взаимодействия с уже имеющимися, с акцентом на поддержку прямых соединений между абонентскими терминалами при наличии сетевого содействия со стороны сотовой инфраструктуры. В данном контексте, важно понять истоки развития данной технологии, а также перспективы ее применения в теории и на практике, которое тесно связано с ожиданием и потребностями абонентов, использующих те или иные услуги прямой связи.

Благодарности

Исследование выполнено при финансовой поддержке РФФ в рамках научного проекта № 16-11-10227.

Литература

1. Jokela T., Ojala J., Olsson T. A Diary Study on Combining Multiple Information Devices in Everyday Activities and Tasks // Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. — ACM, New York, NY, USA, 2015. — Pp. 3903–3912.
2. Pyattaev A., Galinina O., Andreev S., Katz M., Koucheryavy Y. Understanding Practical Limitations of Network Coding for Assisted Proximate Communication // IEEE Journal on Selected Areas in Communications. — 2015. — Vol. 33, no. 2. — Pp. 156–170.
3. Bangerter B., Talwar S., Arefi R., Stewart K. Networks and devices for the 5G era // IEEE Communications Magazine, 2014. — Vol. 52, no. 2. — Pp. 90–96.
4. Vastardis N. and Yang K. Mobile Social Networks: Architectures, Social Properties, and Key Research Challenges // IEEE Communications Surveys & Tutorials. — 2013. — Vol. 15, no. 3. — Pp. 1355–1371.
5. Zhang Y., Pan E., Song L., Saad W., Dawy Z., Han Z. Social Network Aware Device-to-Device Communication in Wireless Networks // IEEE Transactions on Wireless Communications, 2015. — Vol. 14, no. 1. — Pp. 177–190.

UDC 519.254.1

Direct Connectivity of Users in Mobile Social Networks

S. D. Andreev

*Federal Research Center “Computer Science and Control” of RAS
44-2 Vavilova str., Moscow 119333, Russian Federation*

Email: serge.andreev@gmail.com

This paper discusses the history and perspectives in development of mobile social networks, which may rely on the utilization of direct connectivity capabilities between their users.

Key words and phrases: mobile social networks, direct device connectivity.

УДК 621.391

Разработка протокола передачи данных в сетях MANET

А. А. Бахтин

*Кафедра телекоммуникационных систем
Национальный исследовательский университет «МИЭТ»
пл. Шокина, д. 1, Москва, Зеленоград, Россия, 124498*

Email: bah@ieee.ru

Исследование и анализ существующих низкоорбитальных спутниковых систем связи показал, что они обладают рядом существенных недостатков. Следующим этапом эволюционного развития спутниковых систем связи является разработка и построение глобальной интеллектуальной телекоммуникационной спутниковой системы, которая позволит создать универсальную интеллектуальную телекоммуникационную платформу для передачи, обработки, хранения и вычисления данных в орбитальном и наземном сегментах. В работе представлена разработка протокола сетевого уровня децентрализованной самоорганизующейся сети для применения в низкоорбитальных спутниковых системах связи. Предложенный протокол со встроенным алгоритмом маршрутизации, предполагающий возможность поддерживать сетевое взаимодействие между узлами, в качестве которых выступают спутники на низкой орбите, а также роботизированные комплексы в зонах их радиовидимости. Представлена структурная схема обмена данными протокола сетевого уровня с другими уровнями (межуровневое взаимодействие). Приведено математическое обоснование выбора мощности передачи и метрики передаваемого сигнала. Использование выбранной метрики позволяет выбрать маршрут, который учитывает требуемое количество ретрансляций сигнала, позволяет минимизировать расход энергии при передаче, а также поддерживает требуемую вероятность ошибки при передаче сообщения на каждом участке сети.

Ключевые слова: протокол передачи данных, MANET.

1. Разработка протокола передачи данных в сетях MANET

Исследование и анализ существующих низкоорбитальных спутниковых систем связи показал, что они обладают рядом существенных недостатков. Следующим этапом эволюционного развития спутниковых систем связи является разработка и построение глобальной интеллектуальной телекоммуникационной спутниковой системы, которая позволит создать универсальную интеллектуальную телекоммуникационную платформу для передачи, обработки, хранения и вычисления данных в орбитальном и наземном сегментах. Создание данной глобальной спутниковой системы базируется на основе объединения технологий: Интернет, больших данных (Big Data), распределенного хранения и вычисления; программно-определяемого радио (SDR) и методов определения маршрутов передачи больших объемов данных.

Каждая из перечисленных технологий в настоящее время используется в спутниковых системах либо частично, либо не используется вообще. Объединение данных технологий является сложной, и в тоже время, одной из перспективных и стратегически важных задач, отвечающим национальным интересам России [1].

В работе представлена разработка протокола сетевого уровня децентрализованной самоорганизующейся сети для применения в низкоорбитальных спутниковых системах связи. Предложенный протокол со встроенным алгоритмом маршрутизации, предполагающий возможность поддерживать сетевое взаимодействие между узлами, в качестве которых выступают спутники на низкой орбите, а также роботизированные комплексы в зонах их радиовидимости.

Алгоритм маршрутизации децентрализованной самоорганизующейся сети представляет собой составную часть протокола передачи данных сетевого уровня. Протокол сетевого уровня решает набор следующих задач: инкапсуляцию и декапсуляцию данных верхнего или нижнего уровня, логическую адресацию узлов в масштабе сети и количественную оценку при передаче данных на основе метрики. Метрика

выбирается в зависимости от требований, предъявляемых к узлам всей сети, и представляет собой параметры обслуживания, соотношение сигнал/шум, BER, мощность принимаемого/излучаемого сигнала, задержки передачи данных, количество транзитных узлов и т.д.

Структурная схема обмена данными протокола сетевого уровня с другими уровнями [2] представлена на рис. 1.



Рис. 1. Структурная схема обмена данными протокола сетевого уровня

Данный протокол представляет механизм сетевого взаимодействия, обеспечивающий работу узлов в сети с минимальными затратами энергии на передачу.

С учетом помеховой обстановки в спутниковых каналах с некоторым запасом энергии ΔP_{TX} справедливо следующее выражение:

$$P_{TXAxy} = P_{Ny} \cdot (SNR_A - 1) \cdot P_{TXyx} / (P_{RXyx} + P_{Nx}) + \Delta P_{TX}. \quad (1)$$

Мощность передачи сигнала тестового сообщения P_{TXyx} с которой узел y не только оповещает соседние узлы о себе, но и позволяет каждому из них получить оценку оптимальной скорости передачи и минимально необходимую мощность для передачи сообщений до узла y .

Следовательно, в соответствии с выражением (1) энергозатраты минимизируются. С учетом количества ретрансляций метрику маршрутизации можно представить в виде выражения:

$$M = \sum_{i=2}^n t_{TXi} \cdot (P_{Ni-1} \cdot (SNR_{Ai} - 1) \cdot P_{TXi-1} / (P_{RXi} + P_{Ni}) + \Delta P_{TX}). \quad (2)$$

Данное значение метрик (2) позволяет построить маршрут в сети с наименьшими энергетическими затратами.

2. Заключение

В работе предложена структурная схема обмена данными протокола сетевого уровня для низкоорбитальной спутниковой системы связи. Использование выбранной метрики позволяет выбрать маршрут, который учитывает требуемое количество ретрансляций сигнала, позволяет минимизировать расход энергии при передаче, а также поддерживает требуемую вероятность ошибки при передаче сообщения на каждом участке сети.

Литература

1. Бахтин А. А., Волков А. С., Баскаков А. Е. Исследование особенностей реализации алгоритмов доступа к среде в мобильных самоорганизующихся сетях связи // Труды МАИ. — 2017. — № 97. — 18 с. Электронный ресурс: <http://trudymai.ru/published.php?ID=87331>.
2. Бахтин А. А., Волков А. С., Миронов А. В., Муратчаев С. С. Разработка алгоритма маршрутизации в децентрализованных самоорганизующихся сетях MANET // Естественные и технические науки. — М.: ООО Издательство «Спутник+», 2016. — № 12 (102). — С. 232–243.

UDC 621.391

Development of Data Transfer Protocol in MANET

A. A. Bakhtin

*Department of Telecommunications
National Research University of Electronic Technology
1 Shokin Square Bld., Moscow, Zelenograd 124498, Russian Federation*

Email: bah@miee.ru

Research and analysis of existing low-orbit satellite communication systems has shown that they have a number of significant shortcomings. The next stage in the evolution of satellite communication systems is the development and construction of a global intelligent telecommunications satellite system that will create a universal intelligent telecommunications platform for data transmission, processing, storage and computation in the orbital and terrestrial segments. The paper presents the development of a network layer protocol for a decentralized self-organizing network for use in low-orbit satellite communication systems. The proposed protocol with an integrated routing algorithm, which presupposes the ability to maintain network interaction between nodes, which are satellites in low orbit, as well as robotic complexes in their radio visibility zones. A block diagram of the network layer protocol data exchange with other layers is presented (inter-layer interaction). The mathematical justification of the choice of transmission power and the metric of the transmitted signal is given. Using the selected metric allows you to select a route that takes into account the required number of signal retransmissions, allows you to minimize power consumption during transmission, and also supports the required error probability when transmitting a message on each section of the network.

Key words and phrases: data transfer protocol, MANET.

УДК 519.872

Численный анализ простейшей модели транспортного потока в умном городе

Е. Ю. Богданова

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: elenabogdanova1@gmail.com

В докладе приводится постановка задачи управления транспортными потоками. Показана математическая модель движения автотранспорта через регулируемый перекресток, основанная на теории массового обслуживания.

Ключевые слова: анализ инфокоммуникационных систем, умный город, транспортный поток, транспортная сеть, моделирование транспортных потоков.

1. Введение

В настоящее время стремительный рост автомобилей приводит к падению пропускной способности сети автомобильных дорог. Для многих крупных городов данная проблема является одной из центральных, поэтому требует неотложного решения [1]. Основными причинами автомобильных заторов являются неправильный режим работы светофоров, неудобные развязки, а также неблагоприятные погодные условия, которые затрудняют управление транспортным средством. Благодаря транспортному моделированию возможно рациональное управление городской системой, строительство новых и реконструкция уже существующих дорог. Для построения математической модели регулируемого перекрестка применяется теория систем массового обслуживания. Уже на протяжении полувека исследуется автомобильное движение, однако эти исследования не часто применяются на практике из-за нестабильности, многообразия транспортного потока и необходимости ввода все большего количества параметров [2]. Для получения оптимального решения транспортных задач требуется специализированное оборудование и программное обеспечение, что затрудняет оперативное внедрение таких систем в России.

Таким образом, задача актуальна в рамках известной концепции умного города [4]. Понятие «умная» транспортная система заключается в том, что система состоит из отдельно смоделированных участков, которые затем составляются и вносятся в единую базу знаний. За счет постоянного мониторинга и поступающей информации от служб дорожного управления происходит периодическая корректировка информационно-математической модели и повышается уровень предоставления услуг в городах. Моделирование позволяет оценить эффективность управления транспортной сетью города, выявить потенциально проблемные участки для их оперативного устранения [3]. В докладе дается краткий обзор метода и численный анализ, которые затем будут использованы в дальнейших исследованиях по оптимизации транспортных потоков в умном городе.

2. Основная часть

В докладе рассматривается модель, описывающая проезд транспортных средств через регулируемый перекресток, как система массового обслуживания с переменной интенсивностью обслуживания и ограниченной очередью. Основными параметрами исследуемой системы являются: T – длина полного цикла светофора (длина «зеленой» фазы $\tau = \frac{T}{2}$); λ – интенсивность входного потока; μ_0 – интенсивность потока обслуживания в «зеленой» фазе; n – максимальная длина очереди.

Интенсивность обслуживания $\mu(t)$ имеет, например, периодический вид

$$\mu(t) = \begin{cases} \mu_0, & t - [\frac{t}{T}]T < \tau, \\ 0, & \text{иначе.} \end{cases}$$

В докладе приведен пример численного анализа. Вычислена средняя длина очереди машин на перекрестке $r(t) = \sum kp_{k+1}(t)$, $k = 1, \dots, n$, для чего была решена система дифференциальных уравнений Колмогорова и найдено стационарное распределение.

3. Заключение

Большую роль в улучшении транспортной ситуации играет моделирование дорожного движения. В докладе представлены результаты наблюдения за распределением времени пересечения перекрестка автомобилями в зависимости от времени, прошедшего от начала включения разрешающего сигнала, с помощью методов математического моделирования. В дальнейшем планируется исследование моделей, которые позволяют создавать маршруты с учетом указания мест остановок, обгона и реализовывать систему обработки данных и согласование расписаний.

Литература

1. Ахмадинуров М. М., Завалицин Д. С., Тимофеева Г. А. Математические модели управления транспортными потоками: монография. — Екатеринбург: Изд-во УрГУПС, 2011. — 120 с.
2. Вентцель Е. С., Овчаров Л. А. Теория случайных процессов и ее инженерные приложения. Учеб. пособие для вузов. 2-е изд., стер. — М.: Высш. шк., 2000. — 383 с.: ил.
3. Саломатина Е. В. Ресурсный подход к понятию «умный город» // Доклады ТУСУРа. — Томск, 2016. — Т. 19. — С. 137.
4. Shaping smarter and more sustainable cities. ITU-T's Technical Reports and Specification / ITU-T — Geneva, 2016. — 1118 p.

UDC 519.872

Numerical Analysis of the Simplest Model of Traffic Flow in a Smart City

E. Y. Bogdanova

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: elenabogdanova1@gmail.com

The task formulation of managing traffic flows is presented in the report. A mathematical model of the vehicle movement through an adjustable intersection based on the theory of mass service is shown.

Key words and phrases: infocommunication system analysis, smart city, traffic flow, transport network, traffic flow modeling.

УДК 621.391

Разработка двуххранговой помехоустойчивой системы радиосвязи на основе кодового разделения каналов

А. С. Волков, А. В. Солодков, А. Е. Баскаков, С. С. Муратчаев

*Кафедра телекоммуникационных систем
Национальный исследовательский университет «МИЭТ»
пл. Шокина, д. 1, Москва, Зеленоград, Россия, 124498*

Email: leshvol@mail.ru, solodkov_aw@mail.ru, ab.zhelpl@ya.ru, said.muratchaev@gmail.com

В работе рассмотрена организация сотовой сети без базовых станций с 31 активным абонентом на каждую соту. При выделении 4 ШПС сигналов (длиной в 1023 чипа) на передачу для каждого активного абонента реализуется скорость передачи данных $R_c = 3,6$ кбит/с. При этом внутри соты реализуется режим «каждый со всеми» (полносвязанная сеть). Для обеспечения межстанционного взаимодействия организуется временной TDMA кадр с семью временными интервалами и выбирается ансамбль ШПС с длиной 1023. Между сотами также реализуется режим «каждый со всеми» (полносвязанная сеть). При таком режиме работы каждая станция привязки в отведенном ей временном интервале передает шести ближайшим станциям привязки информации от 24 абонентов своей соты, то есть 24 абонента из 31 могут участвовать в межстанционной связи (или передаче данных). Остальные семь физических абонентов каждой передающей соты организуют связь внутри соты (с использованием своей станции привязки для организации TDMA кадра внутри соты). Предлагаемая сотовая сеть без использования базовых станций может быть реализована с применением ШПС-сигналов длиной 2047 чипов, что позволяет значительно улучшить характеристики сети при необходимости.

Ключевые слова: сети связи, ШПС сигналы, межстанционное взаимодействие.

1. Межстанционное взаимодействие

Организация межстанционного взаимодействия рассматривается при плотнейшем расположении 7 равноудаленных станций привязки на двумерной плоскости в виде центральной станции и 6 станций вокруг нее. В качестве исходных данных принимается скорость абонента 3,6 кбит/с, длина ШПС $D_{10} = 1023$, объем кода $V = 1024$, максимальное минимальное расстояние по всему ансамблю. Для организации взаимодействия между станциями необходимо образовать кадровую передачу данных, с 7 временными интервалами (ВИ) на кадр.

Каждому абоненту сети выделяется ВИ на передачу, в остальные ВИ выполняется прием от остальных абонентов сети, то есть организуется полносвязанная сеть. Синхронизация обеспечена использованием технологий GPS или GLONASS.

В таких условиях целесообразен радиальный режим передачи. Одна станция, в выделенный ей ВИ излучает 96 ШПС-сигналов 6 ближайшим станциям, а в остальные ВИ производит пассивную ретрансляцию принятых от соседних станций ШПС-сигналов в выбранных направлениях (возможно применение направленных антенн), на прием для станции выделяется по 128 различных ШПС-сигналов с длиной $D_{10} = 1023$ чипов из общего объема (1024 ШПС-сигналов). В таких условиях скорость передачи выбирается равной 3,6 кбит/с. Для обеспечения выбранной скорости передачи данных производится разделение 16 ШПС-сигналов, выделяемых для передачи информации к одной станции привязки в течении ВИ, на 16 групп по 8 ШПС-сигналов. Следовательно, количество информации на ШПС-сигнал, равно $\log_2 8 \cdot r_c = 3 \cdot r_c$ бит, где r_c — относительная скорость передачи ПК. Таким образом, скорость передачи для физического абонента:

$$R_{аб} = \frac{3 \cdot 4 \cdot r_c}{T_{кадр}} = 3,6 \text{ кбит/с}, \quad R_{ст} = \frac{3 \cdot 32 \cdot r_c}{T_{кадр}} = 8 \cdot R_{аб} = 28,8 \text{ кбит/с}.$$

В межстанционном взаимодействии участвуют $4_{аб} \cdot 6_{ст} = 24$ из 31 абонента в соте. Поскольку станция излучает $16 \cdot 6 = 96$ ШПС-сигналов в ВИ другим станциям, выбранный ансамбль ШПС должен обеспечить разделение ШПС-сигналов в точках приема.

2. Помехозащищенность системы межстанционного взаимодействия

В качестве исходных данных приняты $R_{аб}$ и $R_{ст}$. При $R_c = 1/2$, выбирается код $RS(N, K, D_x)_{GF(2^6)} = RS(64, 32, 33)_{GF(2^6)}$, $D_x = 33 = 2 \cdot T + 1$, $T = 16$. Используя формулу для вычисления вероятности битовой ошибки алгебраических блочных недвоичных кодов и таблицы [1] при $q_{ви} = 10^{-6}$, для указанного кода RS $q_c = 8,1 \cdot 10^{-2}$. Далее $q_{CRS} = 2 \cdot Q_{ОШ}$ расширяющего кода. Тогда для расширяющего кода вероятность блоковой ошибки:

$$Q_{ОШ} = \frac{q_{CRS}}{2} \leq N_d \left(1 - \left(\sqrt{\frac{2E_{ch}}{J_0} d_x} \right) \right).$$

Число ближайших кодовых комбинаций к переданной $N_d = (15 + 95) = 110$. Введем обозначение:

$$\sqrt{\frac{2P_c \cdot \tau_{ch}}{J_0} d_x} = x = \sqrt{\frac{2P_c \cdot 1,2}{J_0 \cdot W} d_x} = \sqrt{\frac{2,4d_x}{J_0 \cdot W/P_c}}, \text{ при } W = \frac{1,2}{\tau_{ch}}.$$

При выравненных входных мощностях P_c :

$$ПЗ = \frac{J_0 W}{P_c} = \frac{P_c(N-1)}{P_c} = (N-1) = (96-1) = 95 = 19,78 \text{ дБ}$$

для сигналов ШПС, что согласуется с нижней границей $ПЗ_{max} = 107,8 = 20,32$ дБ для выбранной длины ШПС-сигналов, определяемой по методике из [2], значит при достоверности $q_{ви} = 10^{-6}$ возможно одновременно передавать не менее 96 ШПС-сигналов.

3. Заключение

Рассмотрено построение сотовой сети без использования базовых станций. В предлагаемом решении выделяется по 4 ШПС-сигнала ($D_{10} = 1023$) на каждого из 31 активного абонента в соте, обеспечивается скорость абонента $R_{аб} = 3,6$ кбит/с, достигается помехозащищенность 19,78 дБ. Рассмотренная сеть позволяет использование ШПС-сигналов с $D_{11} = 2047$ чипов.

Литература

1. Корн Г., Корн Т. Справочник по математике для инженеров и научных работников. — М.: Изд. Наука. Главная редакция физико-математической литературы, 1974. — 832 с.
2. Viterbi A. J. CDMA: principles of spread spectrum communication.- Addison-Wesley Publishing Company, 1995. — 245 p.

UDC 621.391

Development of a Two-Level Anti-Jam Radio System Based on DSSS

A. S. Volkov, A. V. Solodkov, A. E. Baskakov, S. S. Muratchaev

*Department of Telecommunications
National Research University of Electronic Technology
1 Shokin Square Bld., Moscow, Zelenograd 124498, Russian Federation*

Email: leshvol@mail.ru, solodkov_aw@mail.ru, ab.zhe1p@ya.ru, said.muratchaev@gmail.com

The paper considers the organization of a cellular network without using base stations with 31 active user per cell. When 4 spreading signatures per user are selected (1023 chips in length), the data transfer rate is realized for each active user. In this case, a "everyone with everyone" mode (a fully connected network) is realized within the cell. To provide inter-station interaction, a TDMA frame with seven time slots is organized and a set of spreading sequences with a length of 1023 is selected. Between the cells, the "everyone with everyone" mode is also implemented (a fully-connected network). In this mode of operation, each binding station in the time slot allocated to it transmits information from 24 users of its cell to the six nearest binding stations, that is, 24 users out of 31 can participate in inter-station communication (or data transmission). The remaining seven physical users of each transmitting cell organize communication within the cell (using their binding station to organize a TDMA frame within the cell). The proposed cellular network without the use of base stations can be implemented with the use of spread spectrum signals with a length of 2047 chips, which allows to significantly improve the network characteristics, if necessary.

Key words and phrases: communication networks, DSSS, inter-station communication.

УДК 621.391

Исследование и выбор помехоустойчивых кодов в соответствии с требованиями сетей мобильной связи пятого поколения

А. С. Волков, А. В. Лазарев

*Кафедра телекоммуникационных систем
Национальный исследовательский университет «Московский институт
электронной техники»
пл. Шокина, д. 1, Зеленоград, Москва, Россия, 124498*

Email: leshvol@mail.ru, googlaz1@gmail.com

В работе проведено исследование помехоустойчивых кодов по требованиям систем связи пятого поколения (5G). Представлены результаты моделирования и проведена оценка помехоустойчивости выбранных кодов для 5G при фиксированных параметрах — турбокод, LDPC код и полярный код. Для обеспечения высоких скоростей передачи данных требуется снижение энергии, затрачиваемой на передачу одного бита. Турбокоды, LDPC коды и ПК являются конкурирующими схемами кодирования, рассматриваемыми в качестве кандидатов для применения в стандарте 5G. Турбокоды и LDPC конкурируют в различных применениях, поскольку эти коды демонстрируют хорошую производительность. Турбокоды обычно имеют низкую сложность кодирования и высокую степень декодирования, тогда как LDPC коды имеют высокую сложность кодирования, но относительно низкую сложность декодирования. Однако исследователями были открыты многочисленные классы LDPC кодов, некоторые из которых обладают особой структурой, позволяющие реализовать вычислительно эффективные алгоритмы кодирования. Полярные коды имеют два существенных недостатка: низкую эффективность для коротких кодов и кодов средней длины по сравнению с LDPC кодами аналогичной длины и используемый алгоритм последовательного исключения (SC), что влечет за собой низкую пропускную способность. Таким образом, в работе проведен анализ параметров выбранных схем кодирования для достижения характеристик системы 5G.

Ключевые слова: помехоустойчивое кодирование, 5G, LDPC коды, турбокоды, полярные коды.

1. Введение

За последние несколько десятилетий появилось несколько подобных типов кодов - турбокоды, принятые в стандартах мобильной широкополосной связи 3G и 4G, коды с низкой плотностью проверок на четность (LDPC коды), принятые в стандартах WiFi, 10GBase-T, 802.3av и др., а также полярные коды (ПК), появившиеся относительно недавно и все еще плохо изученные. Турбокоды и LDPC коды используют процесс итеративного декодирования, ПК используют последовательный процесс декодирования. Поскольку каналный декодер должен преодолевать неопределенность, возникающую из-за шума, помех и низкой мощности сигнала на фоне шумов и помех, он обычно имеет гораздо большую сложность, чем каналный кодер. В связи с этим именно каналный декодер обычно является основной проблемой при выборе схемы кодирования.

2. Основная часть

В настоящее время известно множество эффективных помехоустойчивых кодов. Однако требования к устройствам пятого поколения мобильной связи существенно сузят возможности по их применению.

При моделировании были использованы конструкции и параметры турбокода из источника [1], LDPC кода, представленного в [2] и полярного кода (см. табл. 1). Канал связи - AWGN, модуляция - BPSK.

Таблица 1

Параметры исследуемых кодов

Параметр	Турбокод	LDPC	Полярный код
K, бит	6144	8100	1024
R	1/3	1/2	1/2
i	6	6	1
Алгоритм декодирования	MAX-LOG-MAP	Min-Sum	FAST-SSC
L, мкс	1660	3298	865
Th, Мбит/с	3789	803.2	190.7

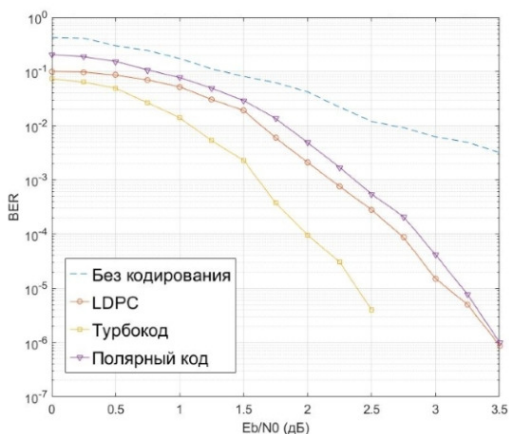


Рис. 1. Сравнение корректирующей способности исследуемых кодов

Как видно из рис. 1, при длинах кодового блока в пределах 8000 бит, турбокоды имеют наилучшую корректирующую способность. Однако LDPC коды демонстрируют относительно хорошие показатели при различных значениях R и K. Также LDPC коды могут показать существенный прирост производительности, путем модифицирования алгоритмов декодирования, проработкой межсоединений элементов декодера и методов синтеза. Полярный код имеет худшую корректирующую способность по сравнению с LDPC кодом и турбокодом, однако ее существенно можно повысить, применяя к алгоритму декодирования циклический избыточный код (CRC). Для достижения оптимальной работы полярных кодов процедура кодирования должна основываться на канале связи. Следовательно, полярные коды не универсальны. Турбодекодер достаточно гибок, так как его структура включает в себя несколько однотипных графов кодовых метрик длины K. Универсальность алгоритма декодирования через такие графы дает возможность изменять скорость кодирования. Турбокоды 4G LTE используют идентичность структурных элементов декодера, что позволяет поддерживать 188 различных кодовых длин K в диапазоне от 40 до 6144 бит.

3. Заключение

В данной работе были исследованы помехоустойчивые коды, предлагаемые для использования в мобильной сети по требованиям пятого поколения 5G. В результате исследования свойств данных схем кодирования было выявлено, что LDPC код может обеспечить достаточную гибкость с помощью различных техник синтеза кода. В отличие от турбокодов, пропускная способность и энергоэффективность LDPC кодов снижается пропорционально скорости кодирования, а задержка масштабируется обратно пропорционально. Турбокоды и LDPC коды обеспечивают требуемую пропускную способность при должной реализации параллельной архитектуры, тогда как для ПК необходимо использовать конвейеризацию. Также турбокоды обладают высокой гибкостью, позволяя подстраивать параметры кода для конкретных задач. Но в то же время LDPC декодеры более просты при аппаратной реализации и при должном синтезе кода не сильно уступают в корректирующей способности турбокодам.

Литература

1. Le Gal B., Jégo C. High-Throughput Multi-Core LDPC Decoders Based on x86 Processor, in: IEEE Transactions on Parallel and Distributed Systems, May 2016.
2. Leonardon M., Cassagne A., Leroux C., Jégo C., Hamelin L-P. and Savaria Y., Fast and Flexible Software Polar List Decoders, Springer Journal of Signal Processing Systems (JSPS), submitted in October 2017.

UDC 621.391

Research and Selection of Channel Codes in Accordance with the Requirements of Mobile Networks of the Fifth Generation

V. S. Volkov, A. V. Lazarev

*Department Telecommunication Systems
National Research University of Electronic Technology
1, Shokin Square, Zelenograd, Moscow, 124498, Russian Federation*

Email: lshvol@mail.ru, googlaz1@gmail.com

The study of error correcting codes on the requirements of communication systems of the fifth generation (5G) was conducted. The simulation results are presented and an estimation of noise immunity of the selected codes for 5G with fixed parameters - turbo code, LDPC code and polar code. To ensure high data transfer rates, the energy required to transmit a single bit is required. Turbo codes, LDPC codes and PCs are competing coding schemes that are considered candidates for use in the 5G standard. Turbo codes and LDPC compete in various applications, because these codes demonstrate good performance. Turbo codes usually have low coding complexity and a high degree of decoding, whereas LDPC codes have high coding complexity, but relatively low decoding complexity. However, researchers have discovered numerous classes of LDPC codes, some of which have a special structure that allow implementing computationally efficient coding algorithms. Polar codes have two significant disadvantages: low efficiency for short codes and medium-length codes compared to LDPC codes of the same length and the algorithm used for sequential elimination (SC), which entails a low throughput. Thus, the analysis of the parameters of the selected coding schemes was performed in order to achieve the characteristics of the 5G system.

Key words and phrases: channel coding, 5G, LDPC codes, turbo codes, polar codes.

УДК 519.217

Реализация модели случайного блуждания для описания движения объектов в ограниченном пространстве

А.С. Вольнец, А.А. Гарибян

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: volynetc.sp@yandex.ru, armen.zed@yandex.ru

Представлено описание симулятора дискретной модели движения объектов, реализующего перемещения абонентов в 2d пространстве для вероятностной модели случайного блуждания.

Ключевые слова: модель случайного блуждания, симуляция, 2d модель.

1. Введение

Моделирование движения различных объектов приобретает большую актуальность в условиях растущего темпа современной жизни. Одной из областей применения таких моделей стали сети подвижной связи, где при оценке качества беспроводного соединения необходимо учитывать перемещения абонентов [1].

Выбор модели движения объектов напрямую зависит от поставленной задачи и способа задания параметров описательной модели. В работе для описания движения объектов выбрана модель случайного блуждания Random Waypoint, в которой для дискретного времени на каждом шаге задаются случайные скорость и направление движения объекта. Также модель учитывает возможность остановки объекта в конце шага на случайное время. В нашем случае будет рассматриваться дискретная модель движения на ограниченном участке 2d пространства и вероятностно заданными условиями изменения направления. В начале шага для каждого независимого объекта разыгрывается скорость, направление и параметры движения.

2. Описание системы

Система состоит из следующих компонентов: 2d область моделирования, N независимых объектов (объекты представляются в виде материальной точки, на первоначальном этапе исключается возможность взаимного влияния объектов друг на друга). Также для дискретного моделирования задаётся общая продолжительность моделирования (параметр $maxtime$), диапазон изменения скорости (от $speedmin$ до $speedmax$) и направления движения - угол отклонения нового направления движения от предыдущего (от 0 до 2π). Прямоугольная область моделирования ограничена точками (x_{pl}, y_{pl}) , (x_{pr}, y_{pr}) , (x_{pl}, y_{pr}) и (x_{pr}, y_{pl}) .

В момент начала моделирования в сгенерированной области случайным образом размещаются N объектов. В начале такта для каждого объекта независимо разыгрывается скорость движения (равномерно распределенная в заданном диапазоне случайная величина) и угол отклонения направления движения на текущем такте от направления на предыдущем такте. При достижении границы области моделирования траектория движения объекта "зеркально" отражается от границы (под тем же углом). В каждый момент времени фиксируется координата нового положения объекта. Таким образом, по окончании процесса моделирования для каждого объекта получаем траекторию его движения, заданную последовательностью точек.

Блок-схема алгоритма моделирования движения приведена на Рис.1

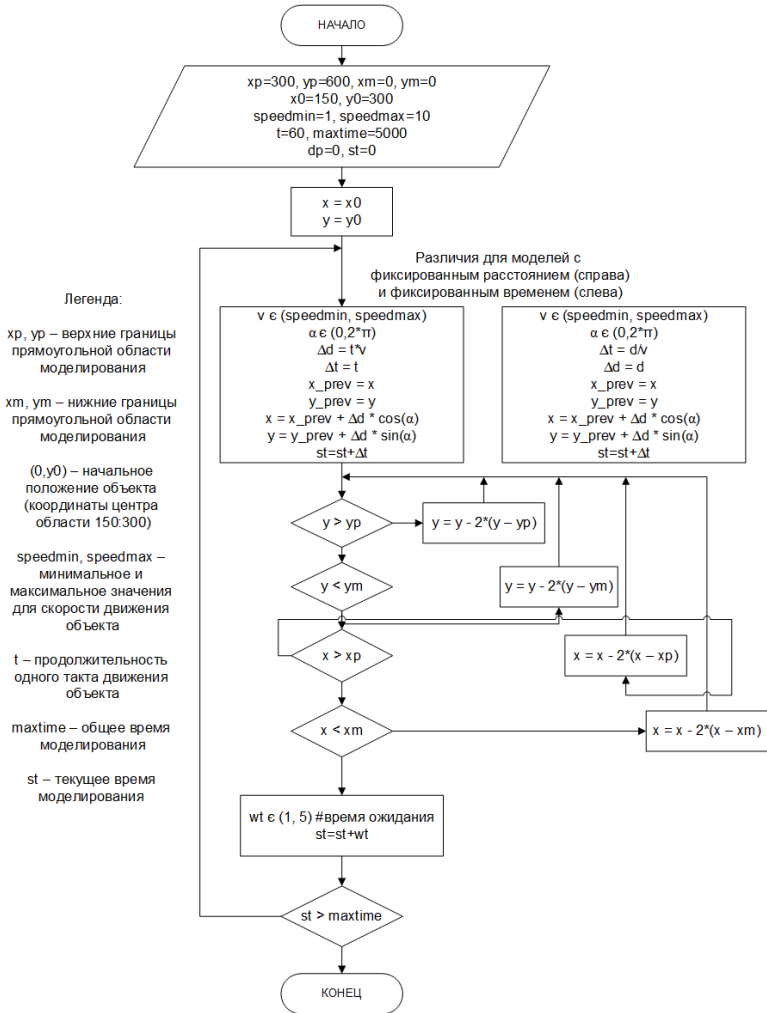


Рис. 1. Блок-схема, реализующая движение объекта с отражением от границ области

Разработанный симулятор может быть использован для анализа интерференции в беспроводной сети связи. В этом случае объектами являются приемно-передающие устройства, а траектории их движения позволяют в каждый момент времени оценить расстояние между объектами, которое существенно влияет на интерференцию [2–4].

На блок-схеме также приведён вариант моделирования движения при котором в начале каждого такта разыгрывается расстояние, которое объект должен пройти за такт. В этом случае длительность такта не фиксирована, а зависит от расстояния и скорости движения на текущем такте.

3. Заключение

Задачей дальнейших исследований является совершенствование разработанного симулятора, добавление в него возможности моделирования пауз в движении, а также учёт препятствий в области моделирования и учёт возможных столкновений объектов. Авторы благодарят профессора Наумова В.А., Исследовательский институт инновации, Хельсинки, Финляндия, за полезные обсуждения на спецсеминаре по моделированию движения.

Литература

1. Tracy Camp, Jeff Boleng, Vanessa Davies, A Survey of Mobility Models for Ad Hoc Network Research, *Wirel. Commun. Mob. Comput.* 2002; 2:483–502 (DOI: 10.1002/wcm.72)
2. Гайдамака Ю.В., Андреев С.Д., Сопин Э.С., Самуйлов К.Е., Шоргин С.Я., Анализ характеристик интерференции в модели взаимодействия устройств с учетом среды распространения сигнала // *Информатика и ее применения.* – 2016. – Т. 10. – Вып. 4. – С. 2–10.
3. Гайдамака Ю.В., Орлов Ю.Н., Молчанов Д.А., Самуйлов К.А., Моделирование отношения сигнал/интерференция в мобильной сети со случайным блужданием взаимодействующих устройств // *Информатика и ее применения.* – 2017. – Т. 11. – Вып. 2. – С. 50–58.
4. Гайдамака Ю.В., Самуйлов К.Е., Шоргин С.Я., Метод моделирования характеристик интерференции при прямом взаимодействии перемещающихся устройств в гетерогенной беспроводной сети пятого поколения // *Информатика и ее применения.* – 2017. – Т. 11. – Вып. 4. – С. 2–9.

UDC 519.217

A Realisation of Random Waypoint Mobility Model Describing the Objects Movement in a Limited Space

A.S. Volynets, A.A. Garibyan

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: volynetc.sp@yandex.ru, armen.zed@yandex.ru

Discret mobility model simulator representing discription of objects movement in 2D space for probabilistic mobility model of random walk.

Key words and phrases: random walk mobility model, simulation, 2d model.

УДК 519.816

Алгоритм принятия решений с помощью вычислений со словами для задач большой размерности

А. А. Гайдамака, Н. В. Чухно, О. В. Чухно

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: aagajdamaka@sci.pfu.edu.ru, nadezdachukhno@yandex.ru, olga-chukhno95@yandex.ru

С каждым днем все больше и больше людей получают доступ в Интернет, где могут общаться, обмениваться информацией и высказывать мнение. Социальные сети становятся не просто площадкой для развлечений, но и мощным механизмом для достижения консенсуса. В докладе представлен новый метод группового принятия решений, который работает в средах, где имеется большое количество экспертов и альтернатив. Метод проиллюстрирован численным примером и показаны результаты работы алгоритма.

Ключевые слова: анализ социальных сетей, нечеткая логика, групповое принятие решений, вычисление со словами, набор языковых выражений.

1. Введение

В настоящее время подавляющее большинство людей в мире могут общаться и обмениваться информацией с помощью сети Интернет [1]. В последние годы исследователи в области принятия решений проявили интерес к социальным сетям, ввиду того, что последние могут быть успешно использованы для моделирования процесса взаимодействия между лицами, принимающими решения [2].

Классический сценарий процесса GDM (англ. group decision-making, GDM) определяет ситуацию, в которой есть некоторая проблема, требующая решения, набор альтернатив и группа лиц, принимающих решения. Последние имеют собственные знания и опыт и передают свои предпочтения или мнения о наборе альтернатив для достижения консенсуса [3]. Цель процесса группового принятия решений состоит в том, чтобы упорядочить различные альтернативы от наилучшей к наихудшей с помощью ассоциации некоторых степеней предпочтений, выраженных в единичном интервале [5]. Процесс GDM - это процесс, ориентированный на людей, с присущей им внутренней неточностью, субъективностью и неопределенностью в выражении оценок. Наиболее подходящим в таких условиях инструментом является теория нечеткой логики [5]. Она используется в течение длительного времени, так как является более богатым и более общим способом представления оценок, чем субъективная вероятность возникновения рассматриваемого события, которое было источником традиционных методов GDM [4].

Тем не менее, необходимо внедрять алгоритмы для принятия решений в группе, которые позволят достаточному числу пользователей предоставлять мнения в большом объеме информации.

2. Метод GDM для большого объема данных

Процесс группового принятия решений включает действия, представленные на рис. 1:

1. Определение количества экспертов и альтернатив.
2. Выбор набора языковых выражений (англ. Linguistic Term Sets, LTS): каждый эксперт выбирает LTS, который лучше соответствует его потребностям.
3. Предоставление предпочтений: эксперты выставляют оценки, используя выбранную LTS. Поскольку количество альтернатив велико, нецелесообразно заставлять каждого эксперта попарно сравнивать всевозможные альтернативы.

4. Преобразование языковых выражений в единый вид: все предпочтения, представленные экспертами, выражаются с использованием единой LTS.
5. Агрегирование: информация объединяется в единую коллективную информационную часть, используя оценки экспертов.
6. Эксплуатация: создается рейтинг альтернатив.
7. Исследование участия экспертов в процессе GDM.

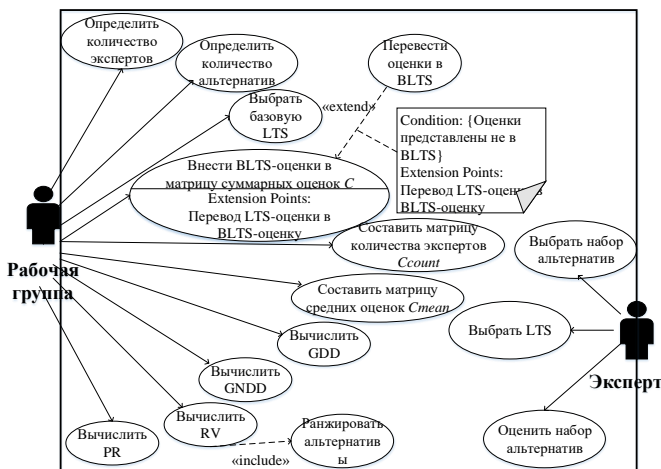


Рис. 1. Use case алгоритма группового принятия решений

Далее представлен пример, иллюстрирующий работу алгоритма процесса GDM.

В результате работы алгоритма создается рейтинг социальных сетей по удобству использования на основе мнений пользователей Интернета. Мы отобрали 27 самых популярных социальных сетей и провели опрос среди 20 экспертов, которые должны были оценивать предложенные варианты, используя попарное сравнение. Процесс GDM подразумевает, что эксперты оценивают не все альтернативы, а лишь выбранный ими набор, о котором они могут судить наверняка. Таким образом, количество значимых предпочтений, которые должен предоставить каждый эксперт, уменьшается. Полный набор альтернатив: x_1 : VK, x_2 : Twitter, x_3 : OK.RU, x_4 : YouTube, x_5 : mail.ru, x_6 : BabyBlog, x_7 : Google Plus, x_8 : WhatsApp, x_9 : Viber, x_{10} : A to Ya, x_{11} : MySpace, x_{12} : LinkedIn, x_{13} : Professional, x_{14} : LiveJournal, x_{15} : Spaces, x_{16} : AlterGeo, x_{17} : Shararam, x_{18} : MySchool, x_{19} : Kindernet, x_{20} : Telegram, x_{21} : Instagram, x_{22} : DeviantArt, x_{23} : Ask.fm, x_{24} : Snapchat, x_{25} : Vine, x_{26} : Tumblr, x_{27} : Facebook. Для обработки результатов опроса была разработана программа. Результаты работы программы можно увидеть в табл. 1.

3. Заключение

Благодаря технологическому прогрессу в мире пользователь сети Интернет становится все более вовлеченным в предоставление информации. Этот факт привел к созданию огромной сети, в которой хранится действительно большой объем данных, доступный каждому пользователю. В докладе представлен новый метод GDM, который способен работать с этими типами сред.

Таблица 1

Ранжирование социальных сетей

x_i	Рейтинг	x_i	Рейтинг	x_i	Рейтинг	x_i	Рейтинг
x_3	73,93	x_4	61,78	x_{11}	43,38	x_{12}	28,50
x_{21}	65,81	x_9	58,56	x_{14}	42,00	x_{24}	27,00
x_8	65,81	x_7	50,83	x_{26}	39,19	x_{16}	25,75
x_5	64,41	x_{10}	50,36	x_{22}	34,57	x_{15}	25,00
x_1	64,14	x_{23}	49,11	x_6	32,62	x_{19}	23,00
x_{20}	63,36	x_{25}	47,19	x_{18}	31,50	x_{13}	21,50
x_{27}	61,87	x_2	46,10	x_{17}	31,00		

Задачей дальнейших исследований является анализ участия экспертов в процессе достижения консенсуса и разработка модели убеждения и аргументации, чтобы иметь дело с лицами, не принимающими решения.

Литература

1. Wasserman S., Faust K. *Social Network Analysis: Methods and Applications*. — Cambridge: Cambridge University Press, 1994.
2. Herrera F., Martínez L. A 2-tuple fuzzy linguistic representation model for computing with words // *IEEE Trans, Fuzzy Syst.* — 2000. — Vol. 8, no. 6. — Pp. 746–752.
3. Pérez I. J., Cabrerizo F. J., Herrera-Viedma E. A mobile decision support system for dynamic group decision-making problems // *IEEE Transactions on Systems, Man, and Cybernetics. Part A: Systems and Humans*. — 2010. — Vol. 40 (6). — Pp. 1244–1256.
4. Herrera-Viedma E., Cabrerizo F., Chiclana F., Wu J., Cobo M., Samuylov K. Consensus in group decision making and social networks // *Studies in informatics and control*. — 2017. — Vol. 26 (3). — Pp. 259–268.
5. Zadeh L. A. Fuzzy sets // *Information and Control*. — 1965. — Vol. 8 (3). — Pp. 338–353.

UDC 519.816

On Group Decision Making Method for Big Data Processing Environment

A. A. Gaidamaka, N. V. Chukhno, O. V. Chukhno

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: aagajdamaka@sci.pfu.edu.ru, nadezdachukhno@yandex.ru, olga-chukhno95@yandex.ru

Every day more and more people get access to the Internet, where they can communicate, exchange information and express opinions. Social networks are not just a place for entertainment, but also a powerful mechanism for reaching a consensus. A new method of group decision making is presented in the report. The algorithm works in environments where there are a large number of experts and alternatives. The example is illustrated and the results of the algorithm are shown.

Key words and phrases: social network analysis, fuzzy logic, group decision making, computing with words, linguistic term sets.

UDC 621.391

Tasks of Ensuring Normalized QoS in Next Generation Networks

B. S. Goldstein*, N. A. Sokolov†

* *Department of Infocommunication Systems*

*The Bonch-Bruевич Saint-Petersburg State University of Telecommunications
22 Prospekt Bolshhevikov, St. Petersburg, 193232, Russian Federation*

† *Department of Network Planning*

*Saint-Petersburg branch of Central science research telecommunication institute
11 Warshavskaya, St. Petersburg, 196128, Russian Federation*

Email: bgold@niits.ru, sokolov@niits.ru

Three tasks related to Quality of Service in Next Generation Communication are discussed. At first glance, the tasks under consideration are not directly related. However, they are united by a common goal. This goal is to provide specified quality of service at various stages of the creation and development of telecommunications networks. The first problem under consideration is related to a qualitative analysis of the errors that arise in the construction of the model of the object or process being studied, the analysis of the chosen model, and the interpretation of the results obtained. The second problem is related to the choice of an adequate distribution for the packet flow at the input of the switching node. It is shown that a great practical interest is associated with the distribution of the dots on a limited time interval. The third task is devoted to the creation of a kind of feedback loop, which will allow analyzing the arising errors during the life cycle of the telecommunications network.

Key words and phrases: Next Generation Communication, Quality of Service, mathematical model, IP-packet.

1. Introduction

The concept of a Next Generation Network (NGN) became the basis for the development of a long-term strategy for the development of information and communication systems for various purposes book [1]. In some instances, Communication Operators and Service Providers modify the NGN concept proposed by the International Telecommunication Union (ITU), but its basic principles remain unchanged. Among these are, in particular, packet-mode technologies for the transmission and switching of all types of the information book [2]. In turn, these technologies stimulated the development and the normalization of new indices of a Quality of Service (QoS). QoS indices for networks of the NGN class are defined in recommendations of the ITU telecommunication standardization sector. The tasks of their ensuring generate a number of complex scientific, technical and organizational problems. Three such problems are discussed in this report. A separate section of the work is devoted to each problem.

2. Features of the study of qualitative characteristics with the help of a mathematical model

QoS characteristics include many metrics and attributes, a significant part of which is defined in the ITU-T E.800 recommendation book [3]. To study the probabilistic-temporal characteristics of QoS, mathematical tools of the telephone traffic theory are usually used. It studies models in the form of telephone traffic systems. For networks of the NGN class, telephone traffic models with a queue and a priority service are usually used.

The use of mathematical models has a number of features which can affect the accuracy of results obtained. Let us consider the algorithm for studying objects or processes performed with the help the model book [4]. It is shown in Figure 1. Let us note that the nature of errors which occur remains the same in the analysis of models which are not typical for the telephone traffic theory.

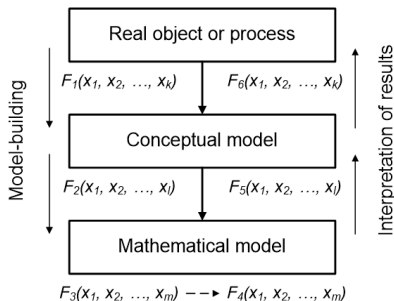


Figure 1. Errors in the study of objects and processes

Let us suppose that a real object or a process can be described by a function $F_1(x_1, x_2, \dots, x_k)$. The set of variables forms a set of parameters which adequately characterize the research subject. When change to the “Conceptual model” block occurs, a number of variables are, as a rule, excluded from the further research for objective and subjective reasons. It means that $l < k$. Of course, the type of the initial function changes too, the record $F_2(x_1, x_2, \dots, x_l)$ reflects. At the “Mathematical model” block, the set of variables under an analysis changes too. Both $l \geq m$ and $l \leq m$ variants are possible. As a result of the study, the function $F_3(x_1, x_2, \dots, x_m)$ is brought to the form, which is designated as $F_4(x_1, x_2, \dots, x_m)$. When interpreting results, a sequential transition to functions $F_5(x_1, x_2, \dots, x_l)$ and $F_6(x_1, x_2, \dots, x_k)$ takes place. The difference in functions $F_1(x_1, x_2, \dots, x_k)$ and $F_6(x_1, x_2, \dots, x_k)$ presents the practical interest, but comparisons of two other pairs of functions also needed for an error analysis. Results of such an analysis make it possible to reveal variables which bring a significant contribution to the accuracy of the model study.

3. The description of an IP packet flow by “heavy tail” distributions

The study of the IP-packet flow showed that it differs distinctly from the Poisson’s arrivals accepted in the telephony. To describe flows of IP packets, distributions with “heavy tails” began to be used book [5]. For such distributions, the coefficient of the variation of intervals between IP-packet arrival times C_A significantly exceeds 1. Measurements carried out in a number of multiservice networks showed that it is also appropriate to use distributions set on a finite time interval for the description of an incoming flow, for which the condition $C_A > 1$ is met too.

In book [6], the typical example is given for the comparison of the average delay time $S^{(1)}$ in a single-line telephone traffic system with a constant service time. At that, two laws of the distribution of the interarrival times of packets: Weibull and Beta. In Kendall’s classification such models are designated as $W|D|1$ and $B|D|1$, respectively. Thus, the distribution with “heavy tail” is chosen for the first model, and the distribution defined on the limited interval is chosen for the second model.

The analysis of functions shows that a substantial delay is inherent in distributions within the limited interval. The behaviours of functions $S^{(1)} = F(\rho)$ are depicted in Figure 2. The value of $S^{(1)}$ is normalized with respect to the average service time $B^{(1)}$. These functions are plotted for the coefficient of variation $C_A = 10$. Delay characteristics were obtained by the simulation.

It is obvious that with equality of the first and second moments of the distribution $A(t)$, delays for distribution defined on the finite interval will be higher than those for

the “heavy tail” distribution. Apparently, this fact served for the selection of the Beta distribution for the study of a machine-to-machine data communication traffic book [7]. In the opinion of authors, similar distributions shall become the target of a research along with fractal processes book [5].

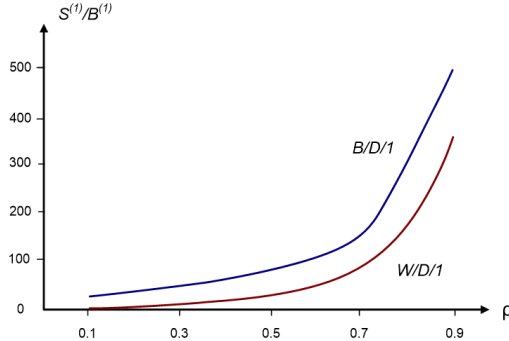


Figure 2. The average delay as the function of a load under $C_A > 1$

4. The support of qualitative indicators by the major players of the market of information and communication services

The analysis of factors leading to a decrease in quality indices of the multiservice traffic service showed that one of main reasons is the lack of an end-to-end monitoring system by the major players of the information and communication market. Figure 3 shows the chain of changes introduced into the hypothetical project of the telecommunication network development book [8]. Such changes are conditioned by a combination of objective and subjective factors, some of which cannot be taken into account beforehand. They are connected in book [8] with tasks of planning networks of the NGN class, but the proposed chain is quite acceptable for studying QoS aspects.

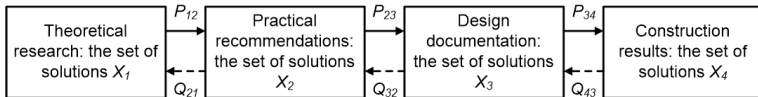


Figure 3. Changes in solutions related to QoS indices

As a result of theoretical studies, a set of solutions is formed for ensuring QoS indices, designated in the considered chain as X_1 . At the stage of compiling practical recommendations in the set of X_1 changes are introduced, which can be considered as distortions. This fact is reflected by the probability P_{12} , which serves as a measure of the difference between X_1 and X_2 sets. The effectiveness of information which is transmitted between the first and second blocks of the model under consideration via the feedback loop is often very low. This effectiveness is estimated by the probability Q_{21} . When developing the design documentation, the set of solutions X_3 is formed. It differs from X_1 and X_2 sets that are estimated by P_{12} and P_{23} probabilities, respectively. The

information effectiveness in the feedback loop between second and third blocks of the proposed model is not so high too. The information effectiveness in the feedback loop between the second and third blocks of the proposed model is also not so high. It is expressed through the probability Q_{32} . Characteristics of the hardware complex to be used by the Communications Operator to support the normalized QoS indices are determined by the X_4 set. It is obvious that its difference from the X_1 set can become significant. Remarks on probabilities P_{34} and Q_{43} coincide with those assertions which were expressed above about similar values differing only in subscripts. An effectively functioning information system with a feedback, similar to the chain under consideration, will allow using modern approaches to solving problems in terms of QoS. For this purpose, modern technologies like Big Data book [9], Data Mining book [10] and neural networks book [11] will be useful.

5. Conclusions

Three tasks devoted to providing normalized quality indices in the next generation communication network do not reflect all aspects of such complicated problem. Nevertheless, they allow identifying a number of important points, which the authors wanted to draw the readers' attention to.

Firstly, the development of mathematical models and their research shall be accompanied by a thorough analysis of the adequacy of obtained results. This concerns even intermediate results, since sources of errors which can appear in the future can be revealed in them. First of all, such danger can arise when changes happen in the behaviour of the IP-packets flow due to the appearance of new services and applications.

Secondly, you should pay careful attention to the expediency of using new (for tried out models) class of incoming IP packet flow distributions. These distributions are defined on a limited time interval, that is, "tails" are non-existent for them at all. On the other hand, high values of the coefficient of variation of the intervals durability between times of the IP packets arrival in the switching node are inherent for distributions under the consideration.

Thirdly, to provide normalized quality indices in the next generation communication network, it is needed to create an effective information exchange system which operates at all stages of the life cycle of the information and communication system and takes into account interests of the major participants of the relevant market. This system should use advanced technologies for collecting and processing information to develop recommendations that are close to optimal solutions.

References

1. Plevyak T., Sahin V. Next Generation Telecommunications Networks, Services, and Management. — Wiley, 2010, 328 p.
2. Mazda F. Switching Systems and Applications (Focal Telecommunications Pocket Book Series). — Focal Press, 1996, 288 p.
3. ITU-T. Definitions of terms related to quality of service. Recommendation E.800. — Geneva, 2008, 30 p.
4. Myshkis A. D. Elements of the theory of mathematical models. — M.: Book House "Libricom", 2011, 192 p.
5. Shelukhin O. I., Tenyakshev A. M., Osin A. V. Fractal processes in telecommunications. — M.: Radiotekhnika, 2003, 480 p.
6. Levakov A. K., Sokolov A. N., Sokolov N. A. Models of incoming traffic in packet networks. — T-Comm, 2015, №5, pp. 91–94.
7. 3GPP. TR 37.868 "Technical Specification Group Radio Access Network; Study on RAN Improvements for Machine-type Communications; Release 11". — Sophia Antipolis, 2011, 28 p.
8. Levakov A. K. Features of the next generation network functioning in emergency situations. — M.: IRIAS, 2012, 108 p.

9. Erl T., Khattak W., Buhler P. Big Data Fundamentals: Concepts, Drivers & Techniques. — Prentice Hall, 2015, 218 p.
10. Han J., Kamber M., Pei J. Data Mining. Concept and Techniques. — Morgan Kaufmann Publishers, 2011, 703 p.
11. Haykin S. Neural Networks And Learning Machines. — Pearson India, 2016, 944 p.

УДК 621.391

Задачи безопасности нормализуемого QoS в сетях следующего поколения

Б. С. Гольдштейн*, Н. А. Соколов†

** Кафедра информационных систем
Санкт-Петербургский государственный университет телекоммуникаций им.
М. А. Бонч-Бруевича*

Проспект большевиков, д. 22, Санкт-Петербург, Россия, 193232

*† Кафедра сетевого планирования
Ленинградское отделение ЦНИИС (ЛО ЦНИИС)
ул. Варшавская, д. 11, Санкт-Петербург, Россия, 196128*

Email: bgold@niits.ru, sokolov@niits.ru

Обсуждаются три задачи, связанные с качеством обслуживания в сетях следующего поколения. На первый взгляд, рассматриваемые задачи напрямую не связаны. Однако их объединяет общая цель. Эта цель заключается в предоставлении определённого качества обслуживания на разных этапах создания и развития телекоммуникационных сетей. Первая рассматриваемая проблема связана с качественным анализом ошибок, возникающих при построении модели исследуемого объекта или процесса, анализа выбранной модели и интерпретации полученных результатов. Вторая проблема связана с выбором адекватного распределения для потока пакетов на входе коммутационного узла. Показано, что большой практический интерес связан с распределением точек на ограниченном временном интервале. Третья задача связана с созданием своего рода обратной связи, которая позволит анализировать возникающие ошибки в течение жизненного цикла телекоммуникационной сети.

Ключевые слова: сети связи следующего поколения, качество обслуживания, математическая модель, IP-пакет.

UDC 004.4

Performance Evaluation of Dynamic LSA Operation Through a Model of a Stand-Alone Cell

I. A. Gudkova*[†], J. A. Romanovskaya*

* *Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

[†] *Institute of Informatics Problems, FRC CSC RAS
44-2 Vavilova str., Moscow 119333, Russian Federation*

Email: igudkova@sci.pfu.edu.ru, romanovskaya-julia2011@yandex.ru

This paper considers an analytical model of an LTE network using LSA concept to gain access to the airport spectrum according to the limit power algorithm with signal-interference ratio threshold as a stand-alone cell during the airplane takeoff.

Key words and phrases: LTE; path loss; LSA; interference; limit power.

1. Introduction

In this paper, we study an LSA use case, where the airport owns a spectrum license over a large area and uses it for the telemetry when airplanes take-off. There is a cellular network present in the area where the airplane receives telemetry signals, and the respective mobile network operator (MNO) has means to constrain its interference towards the airplanes. Also we assume that planes take-off only occasionally, that is, only one airplane is present in the MNO coverage at once. Thus the spectrum is used in small and localized area around the airplane. In the considered scenario, a mobile network uses an airports telemetry spectrum until an airplane needs to receive telemetry signal from the air traffic control. When it happens, the MNO restricts interference its user equipment (UE) causes around the position of the airplane, to let it receive the telemetry signal. To this end, MNO uses limit power policy. [1] The MNO reduces its UE power for users using LSA band in this area.

2. Main section

Statement 1. (interference level) At the time $t > 0$:

1) We consider that the airplane is moving by the uniformly accelerated motion along a certain line/ with certain parameters:

$$\begin{aligned}x(t) &= x_a + \left(\nu_0 t + \frac{at^2}{2} \cos \beta \right) \cos \gamma, \\y(t) &= y_a + \left(\nu_0 t + \frac{at^2}{2} \cos \beta \right) \sin \gamma, \\z(t) &= \frac{at^2}{2} \sin \beta.\end{aligned}$$

2) The distances from the airplane to the airport and to the subscriber are equal, respectively:

$$\begin{aligned}d_a(t) &= \sqrt{(x(t) - x_a)^2 + (y(t) - y_a)^2 + (z(t) - h_a)^2}, \\d_u(t) &= \sqrt{(\sqrt{d_c^2(t) - z^2(t)} - r_c)^2 + (z(t) - h_u)^2}.\end{aligned}$$

3) The power of the signal which is received by the airplane from the airport transmitter and from the subscriber unit (interfering signal) is calculated by formulas:

$$p_a^{rx}(t) = p_a^{tx} - PL(d_a(t)),$$

$$p_u^{rx}(t) = p_u^{tx} - PL(d_u(t)).$$

4) Then the signal-to-interference ratio from the subscriber unit to the airplane is:

$$SIR = \frac{p_a^{rx}(t)}{p_u^{rx}(t)}.$$

Statement 2. (Threshold power of the subscriber unit).

If, at time $t > 0$, the interference $SIR(t)$ from the subscriber unit to the airplane exceeds the threshold value

$$SIR_0 : SIR(t) \geq SIR_0$$

then the transmitter power of the subscriber unit must be reduced to the value:

$$\overline{p_u^{rx}(t)} = \min\{p_a^{rx}(t) - SIR_0 + PL(d_u(t)), p_u^{rx}(t)\}.$$

Statement 3. (The radius of the radio shadow of the airplane) [2].

At the time $t > 0$, the radius of the radio shadow of the airplane is calculated by formula:

$$d_u(t) = \frac{c}{4\pi f} 10^{p_u^{tx} - \frac{p_u^{rx}(t)}{20}}.$$

Statement 4. (The condition that none of the subscriber in the cell is not required to reduce the power).

At time $t > 0$ the distance from the airplane to the nearest subscriber in the cell is calculated by the formula:

$$d_u^*(t) = \sqrt{D_u^{*2}(t) + z^2(t)},$$

$$D_u^*(t) = \begin{cases} 0 & D_c(t) \leq r_c, \\ |D_c(t) - r_c| & D_c(t) > r_c, \end{cases}$$

$$D_c(t) = \sqrt{(x(t) - x_c)^2 + y_c^2}.$$

Statement 5. (The time interval during which it is required to reduce the power for at least one subscriber unit in a cell).

The time interval during which it is required to reduce power for at least one subscriber unit in a cell is determined by the roots of the algebraic equation

$$\left| \sqrt{x_a + (\nu_0 T + \frac{at^2}{2} \cos \beta) \cos \gamma - x_c)^2 + (y_a + (\nu_0 t + \frac{at^2}{2} \cos \beta) \sin \gamma - y_c)^2 - r_c} \right| =$$

$$= \sqrt{-\frac{a}{2} t^4 + \frac{c}{4\pi f} 10^{p_u^{tx} - \frac{p_u^{rx}(t)}{20}}}.$$

1) If there are no real positive roots or there is only one real root, then power reduction is not required.

2) If there are two real roots and at least one of them is positive,

$$t_1 < t_2, \quad t_2 > 0,$$

then the time interval has the form:

$$[t_1^+, t_2],$$

where

$$t_1^+ = \begin{cases} 0 & t_1 \leq 0, \\ t_1 & t_1 > 0. \end{cases}$$

3. Conclusions

In this paper we studied a stand-alone cell scenario for limit power policy with SIR as the licensee QoS parameter for LTE network using LSA. In this paper, we derived formulas that allow us to calculate the level of interference, the threshold power, the time intervals during which it is required to reduce the power. We use limit power policy here, if you only lower the power and not shut down cells. These formulas allow us to evaluate dynamic LSA operation through a model of a stand-alone cell.

Acknowledgments

The publication was prepared with the support of the PFUR's program "5-100" and with the financial support of the Russian Foundation for Basic Research in the framework of scientific projects No. 17-07-00142 and No. 17-07-00845.

References

1. Aleksei Ponomarenko-Timofeev, Alexander Pyattaev, Sergey Andreevy, Yevgeni Koucheryavy, Markus Mueck, Ingolf Karls. Highly Dynamic Spectrum Management within Licensed Shared Access Regulatory Framework. IEEE Communications Magazine 54(3), p.100-109, 2016.
2. Evgeniy Mokrov, Irina Gudkova. Performance evaluation of dynamic LSA operation through a model of a stand-alone cell, p.35-41, 2016 Societies.

УДК 004.4

Оценка производительности работы LSA на примере модели отдельно взятой соты

И. А. Гудкова*[†], Ю. А. Романовская*

* Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198

[†] Институт проблем информатики
Федеральный исследовательский центр «Информатика и управление» РАН
ул. Вавилова, д. 44, к. 2, Москва, Россия, 119333

Email: igudkova@sci.pfu.edu.ru, romanovskaya-julia2011@yandex.ru

В данной статье построена аналитическая модель сети LTE, использующая мощности аэропорта по технологии LSA, согласно алгоритму ограничения мощности с пороговым значением SIR на примере отдельно взятой соты в моменты взлета самолета.

Ключевые слова: LTE; затухание сигнала; LSA; интерференция; предельная мощность.

УДК 519.218.31

Анализ механизмов повышения энергоэффективности облачных систем

А. В. Дараселия

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: nastyadar6@gmail.com

Построены математические модели для различных механизмов повышения энергоэффективности систем облачных вычислений путем перевода серверов в спящий режим.

Ключевые слова: облачные вычисления, энергоэффективность, теория массового обслуживания.

1. Введение

В целях повышения энергоэффективности системы облачных вычислений сервера могут быть переведены в режим ожидания при низкой нагрузке. Перевод в режим ожидания, с одной стороны, позволяет снизить энергопотребление, а с другой стороны, приводит к дополнительным затратам на включение/выключение сервера. Поэтому важно понимать, при каких условиях будет выгодно перевести сервера в режим ожидания, а при каких – оставить их в рабочем состоянии. Существуют различные механизмы перевода сервера в спящий режим, в данной работе проведен анализ нескольких из этих механизмов.

2. Основная часть

Рассматриваются системы массового обслуживания, состоящие из N групп приборов (серверов), каждая из которых состоит из C приборов. Для краткости, в данной работе приведем результаты анализа для одного сервера.

В базовой модели на систему поступает пуассоновский поток заявок с интенсивностью λ , время обслуживания заявок, а также периоды разогрева и выключения распределены экспоненциально с параметрами μ , α и β , соответственно. Если система пуста, то при поступлении заявки начнется процедура разогрева. Это время экспоненциально распределено с параметром α . Если система находится в рабочем режиме, то как только она останется пустой, сразу переходит в режим выключения. Если за время выключения в систему не поступили новые заявки, то система переходит в спящий режим, а если заявки пришли, то в момент выключения сервер сразу переходит в режим включения. С учетом описания введем пространство состояний, которое задается парой (s, k) , где k – количество заявок на сервере, s – состояние сервера, где $s = 0$ в состоянии ожидания, 1 – включения сервера, 2 – работы и 3 – выключения соответственно. Таким образом, множество состояний S_1 представляется в следующем виде: $S_1 = \{(s, k) | s = 1, 2, 1 \leq k \leq C\} \cup \{(s, k) | s = 3, 0 \leq k \leq C\} \cup (0, 0)$.

Во второй модели [3], в отличие от базовой модели, в которой считалось, что сервер выключается сразу же, как только остается пустым, была рассмотрена система, в которой сервер выключается по прошествии некоторого случайного времени после того, как он остался пустым. Для такой системы множество состояний S_2 представляется в следующем виде: $S_2 = \{(s, k) | s = 1, 1 \leq k \leq C\} \cup \{(s, k) | s = 2, 3, 0 \leq k \leq C\} \cup (0, 0)$. Время ожидания поступления новой заявки, в течении которого система не переходит в спящий режим, экспоненциально распределено с интенсивностью γ .

В третьей модели, так же, как и в базовой модели, система переходит в режим выключения сразу после того, как остается пуста, но в режим включения

переходит только тогда, когда после начала поступления заявок прошло некоторое случайное время θ . Для этой системы множество состояний S_3 представляется как $S_3 = \{(s, k) | s = 0, 3, 0 \leq k \leq C\} \cup \{(s, k) | s = 1, 2, 1 \leq k \leq C\}$. В четвертой модели система переходит из спящего режима в режим включения только после поступления некоторого κ колличесива заявок. В этом случае множество состояний S_4 записывается в виде $S_4 = \{(s, k) | s = 0, 0 \leq k \leq \kappa - 1\} \cup \{(s, k) | s = 1, \kappa \leq k \leq C\} \cup \{(s, k) | s = 2, 1 \leq k \leq C\} \cup \{(s, k) | s = 3, 0 \leq k \leq C\}$.

Для каждой модели были выписаны системы уравнений равновесия, которые были решены аналитически и позволили получить стационарные вероятности. Для оценки энергоэффективности системы были введены формулы расчета средней потребляемой сервером мощности и среднее время пребывания заявки в системе, полученное из формулы Литтла. Для выведенных формул был проведен численный анализ и построены графики для значений $C = 7, \mu = 20, \alpha = 1$ и $\beta = 2$.

Для каждой из трех моделей было отмечено, что при $\gamma \rightarrow \infty$ (см. рис. 1), $\theta \rightarrow \infty$ (см. рис. 2) и $\kappa \rightarrow 1$ (см. рис. 3), соответственно, каждая из трех систем стремится к начальной базовой модели. Параметры γ, θ и κ монотонно влияют на параметры энергопотребления, поэтому нужно учитывать время пребывания заявки, как с точки зрения энергопотребления, так и с точки зрения производительности.

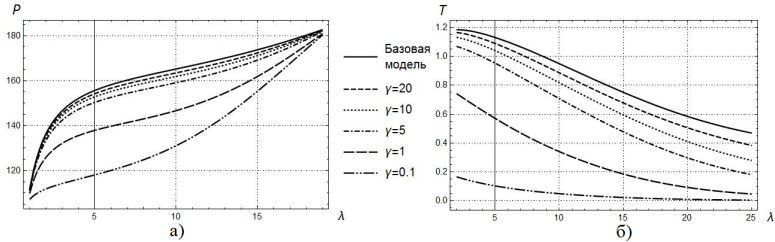


Рис. 1. Графики зависимости: а) мощности P от интенсивности нагрузки λ для разного времени ожидания новой заявки γ ; б) среднего времени пребывания заявки в системе T от интенсивности нагрузки λ

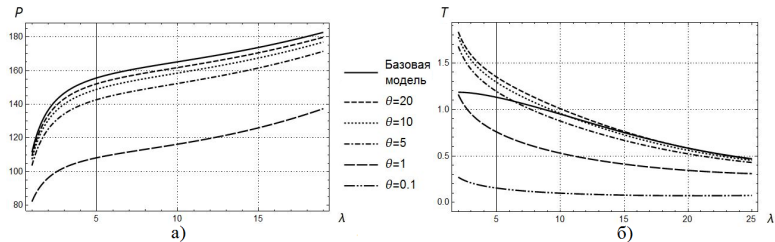


Рис. 2. Графики зависимости: а) мощности P от интенсивности нагрузки λ для разного времени θ ; б) среднего времени пребывания заявки в системе T от интенсивности нагрузки λ

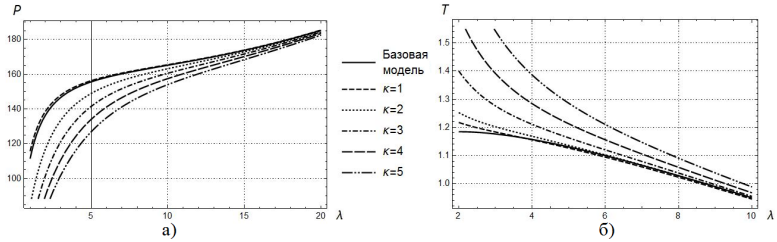


Рис. 3. Графики зависимости: а) мощности P от интенсивности нагрузки λ для разного κ ; б) среднего времени пребывания заявки в системе T от интенсивности нагрузки λ

3. Заключение

В этой работе мы рассмотрели 4 математические модели с различными механизмами повышения энергоэффективности систем облачных вычислений.

Численный анализ показал, что второй механизм дает улучшение и в мощности, и во времени ожидания, для третьего механизма наблюдается улучшение почти на всех промежутках, кроме времени обслуживания при малых нагрузках, четвертый механизм дает улучшение для мощности, но ухудшение по времени.

Литература

1. Conejero J., Rana O., Burnap P., et al. Analysing Hadoop Power Consumption and Impact on Application QoS // Future Generation Computer Systems. — Feb. 2016. — Vol. 55, issue C. — Pp. 213–223.
2. Beloglazov A., Abawajy J., Buyya R. Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing // Future Generation Computer Systems. — 2012. — Vol. 28. — Pp. 755–768.
3. Дараселия А. В., Сопин Э. С. О задаче оптимизации энергопотребления облачной инфраструктуры // XVI Международная конференция имени А. Ф. Терпугова «Информационные технологии и математическое моделирование» (ИТММ–2017), 255–261.

UDC 519.218.31

On the Analysis of Mechanisms for Improving Cloud Computing System Energy Efficiency

A. V. Daraseliya

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: nastyadar6@gmail.com

We consider the mathematical models for various mechanisms to improve energy efficiency of cloud computing systems.

Key words and phrases: cloud computing, energy efficiency, queuing system.

УДК 519.254.1

Об одном алгоритме моделирования движения абонентов в беспроводной сети

А. Н. Жданков*, Ю. В. Гайдамака*[†]

* *Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

[†] *Институт проблем информатики
Федеральный исследовательский центр «Информатика и управление» РАН
ул. Вавилова, д. 44, к. 2, Москва, Россия, 119333*

Email: a.zhdankov@inbox.ru, gaydamaka_yuv@rudn.university

Одной из моделей индивидуальной мобильности объектов является т.н. «модель движения по сетке». Это модель в дискретном времени, когда на каждом временном такте объект передвигается на единичное расстояние с заданной вероятностью. В работе построена модификация этой модели, которая описывает движение объекта к заданной точке, причём перемещение объекта на каждом шаге разыгрывается случайно с учетом направления, определенного заданной конечной точкой.

Ключевые слова: симулятор движения, информационные технологии, отношение сигнал интерференция.

1. Постановка задачи

В работе с помощью модели движения «по сетке» (grid movement) моделируется перемещение n объектов, соответствующих приемо-передающим устройствам беспроводной сети связи. Координаты начальных и конечных точек перемещения каждого объекта разыгрываются с помощью равномерного распределения. После проверки условия на то, что начальное и конечное положение не лежат достаточно близко друг к другу, симулятор переходит к моделированию движения объектов. Объекты движутся по сетке, т.е. каждый объект имеет возможность изменить текущие координаты на единицу в сторону направления движения [1], определенного конечной точкой. Симуляция происходит для заданного числа T шагов. Новые координаты на следующем шаге выбираются с помощью углов α и β , где угол α — угол между вектором направления и осью абсцисс, а β — угол между одним из тех возможных направлений и осью абсцисс, для которого разность между углами α и β минимальна. К примеру, на рисунке 1 показано, как может двигаться абонент при начальной точке А и конечной В. В данном примере, для объекта возможны варианты перемещения на новые точки с координатами (2, 3), (3, 3) или (3, 2). Поскольку минимальным является угол между диагональю и вектором направления, объект передвигается в точку с координатами (3, 3).

Моделирование проводится с целью оценки интерференции в сети беспроводной связи. Два из n объектов соответствуют целевой паре «приемник-передатчик», остальные $n - 2$ соответствуют интерферирующим передатчикам, при этом интерференция на приемнике целевой пары (объект № 1) оценивается по следующей формуле:

$$SIR = S(l)/I(l). \quad (1)$$

Здесь SIR, Signal to interference and noise ratio — показатель, традиционно используемый для оценки качества сигнала в радиоканале, который влияет на скорость передачи данных в канале [2]. В числителе формулы (1) мощность полезного сигнала от целевого передатчика (объект № 2), в знаменателе — суммарная мощность сигналов интерферирующих передатчиков (объекты от №3 до № n). Мощность принимаемого приемником сигнала, как полезного, так и интерферирующего, зависит от расстояния между приемником и передатчиком, базовой мощности g сигнала

передатчика и коэффициента γ затухания в среде распространения радиосигнала, принимающего значения от 2 при прямой видимости до 6 в худшем случае и определяется формулой $g * l^{-\gamma}$.

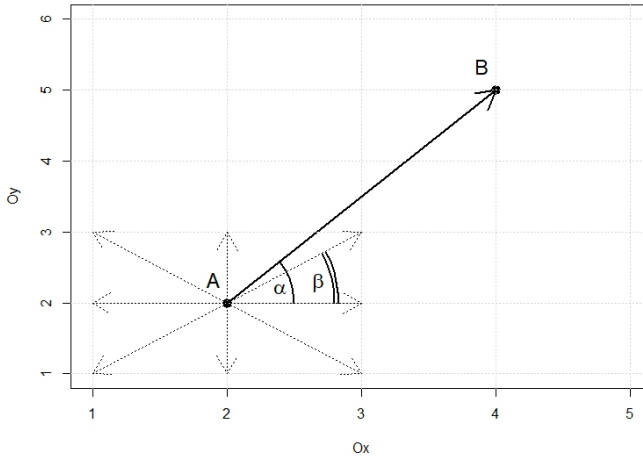


Рис. 1. Выбор направления движения на каждом шаге

Исходные данные: число объектов моделирования $n = 10$, соответственно число интерферирующих передатчиков $n - 2$. Целевая пара приемник-передатчик: первый и второй соответственно. Коэффициенты затухания полезного и интерферирующих сигналов принимают значения $\gamma_1 = 2$ и $\gamma_2 = 3$ соответственно. Область моделирования представляет собой квадрат со стороной $X = Y = 500$ шагов. Расстояние между начальной точкой движения и конечной должно быть не менее, чем 0,1 длины стороны квадрата, иначе координаты для конечной точки разыгрываются снова. Количество шагов моделирования $T = 800$.

Шаг 1. Разыгрываем случайные величины, определяющие координаты каждого из передатчиков: $x_i \sim \text{Random}[\frac{X}{100}, X - \frac{X}{100}]$ и $y_i \sim \text{Random}[\frac{Y}{100}, Y - \frac{Y}{100}]$, $i = 1, \dots, n$.

Шаг 2. Разыгрываем случайные величины, определяющие координаты конечной точки для каждого из передатчиков. $x'_i \sim \text{Random}[\frac{X}{100}, X - \frac{X}{100}]$ и $y'_i \sim \text{Random}[\frac{Y}{100}, Y - \frac{Y}{100}]$, $i = 1, \dots, n$.

Шаг 3. Производим проверку расстояния между начальными и конечными координатами. При выполнении условия повышаем i на единицу, иначе возвращаемся на шаг назад для того же i и находим новые конечные координаты.

Шаг 4. Вычисляем угол α для i -го передатчика.

Шаг 5. Выбираем направление движения из трех возможных.

Шаг 6. Получаем новые координаты i -го передатчика.

Шаг 7. Проверяем, достиг ли передатчик конечной области. Если да, то перезаписываем начальное положение и разыгрываем новую конечную точку.

Шаг 8. Вычисляем полезный сигнал (между целевой парой) и суммарный сигнал от всех интерферирующих передатчиков.

Шаг 9. Находим SIR по формуле (1) для конкретного шага t , $t = 1, \dots, T$.

В результате работы алгоритма получаем значения величины SIR для полученного набора траекторий движения объектов. Анализ собранной статистики позволит определить характеристики случайной величины SIR .

Задачей дальнейших исследований является совершенствование алгоритма моделирования для учета особенностей перемещения абонентов сетей беспроводной связи. Например, для движения внутри помещений необходимо учитывать препятствия, а также особенности движения объекта при приближении к препятствию, которое может отличаться от движения объекта на открытом пространстве.

Благодарности

Исследование выполнено при частичной финансовой поддержке РФФИ в рамках научных проектов № 17-07-00845 и 18-07-00576.

Литература

1. Camp T., Boleng J., Davies V. A survey of mobility models for ad hoc network research // *Wireless Communications and Mobile Computing* *Wirel. Commun. Mob. Comput.* — 2002. — Vol. 483(2). — P. 502. doi:10.1002/wcm.72.
2. Гайдамака Ю. В., Андреев С. Д., Сопин Э. С., Самуйлов К. Е., Шоргин С. Я. Анализ характеристик интерференции в модели взаимодействия устройств с учетом среды распространения сигнала // *Информатика и ее применения.* — 2016. — Т. 10, вып. 4. — С. 2–10.

UDC 519.254.1

About One Algorithm of Motion Simulations of Subscriber in Wireless Network

A. N. Zhdankov*, Yu. V. Gaidamaka*†

* *Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

† *Institute of Informatics Problems, FRC CSC RAS
44-2 Vavilova str., Moscow 119333, Russian Federation*

Email: a.zhdankov@inbox.ru, gaydamaka_yuv@rudn.university

An example of moving of individual mobility is a so-called “grid movement”. This model is carried out in discrete time when the object is moving on unit distance with a probability. In this work a modification of the model which describes the moving of the object to the given point is created. Furthermore, the movement of the object at each step is being played out taking into account a direction that is defined by the given final point.

Key words and phrases: motion simulator, information technology, signal to interference plus noise ratio.

УДК 51-74

Разработка метода выявления аномалий сетевого трафика

Е. В. Карачанская*[†], Н. И. Соседова[†]

* *Кафедра информационных технологий и систем
Дальневосточный государственный университет путей сообщения
ул. Серышева, д. 47, Хабаровск, Россия, 680027*

[†] *Кафедра высшей математики
Тихоокеанский государственный университет
ул. Тихоокеанская, д. 136, Хабаровск, Россия, 680035*

Email: elena_chal@mail.ru, sosedova.nadezhda@yandex.ru

В работе предложен новый метод обнаружения аномалий сетевого трафика, моделируемого фрактальным броуновским движением. Метод относится к группе статистических методов, основан на полуконтролируемой методике обнаружения Алгоритм анализа трафика на аномальность состоит из двух частей: расчёт эталонных значений (обучение) и обнаружение аномалий. Расчёт эталонов строится на значении параметра Хёрста для некоторых показателей из заголовков пакетов. Для уменьшения числа ложных срабатываний используются несколько временных интервалов для наблюдения. Работа метода сводится к сравнению показателей на этапе расчета эталонов с показателями, рассчитываемыми в процессе мониторинга аномалий. Данный метод может быть использован для мониторинга как аномалий во входящем, так и в исходящем трафике, что позволяет его включить как в комплекс обнаружения вторжений, так и в DLP-системы.

Ключевые слова: аномалии сетевого трафика, фрактальное броуновское движение, показатель Хёрста, RS-анализ.

1. Введение

Вопрос обнаружения аномалий сетевого трафика (АСТ), обусловленных внешним вторжением, уже достаточно долгое время не покидает рейтинга наиболее актуальных проблем в области защиты информации. При этом зачастую существующие методы выявления АСТ, используемые в системах обнаружения вторжений, если и являются эффективными, то достаточно сложны в программной реализации.

Под *сетевым трафиком* будем понимать объём информации, переданный по сети за определённое время. Под *аномалиями сетевого трафика* будем понимать любые отклонения показателей сети от заранее зафиксированных в качестве эталонных.

В общем случае, обнаружение АСТ позволяет выявлять атаки на сеть, следить за отклонениями канала, отслеживать работу сотрудников и т.п.

Исследования свойств сетевого трафика показывают, что они значительно отличаются от свойств трафика телефонных сетей и имеют фрактальные свойства [2–4]. Наличие данного свойства у сетевого трафика означает, что простые модели, использующие пуассоновское распределение, неточны, и сети, построенные без учёта самоподобия, могут функционировать в непредсказуемых режимах. Фрактальное броуновское движение было предложено как модель для «свободного трафика», агрегированного от множества независимых источников. Модель привлекательна за счёт простоты анализа и за счёт того, что она характеризуется долгосрочной зависимостью. Существующие методы выявления АСТ обладают определёнными недостатками. Наиболее оптимальным, возможно, является метод, предложенный Д. Унтеровым [1], за исключением недостаточного размера выборки для использования параметра самоподобия и сложностью аппаратной реализации. Данный метод был взят за основу для создания нового, но был произведен пересмотр необходимого времени наблюдения для корректного учёта свойства фрактальности сетевого трафика, внесены изменения в порядок расчёта и аппаратная реализация была заменена программной, что в свою очередь позволяет использовать данный метод как модуль и комбинировать его с другими методами для большей эффективности.

2. Выявление аномалий сетевого трафика

Фрактальное броуновское движение (ФБД) — случайный процесс, который обладает некой памятью. Для аппроксимации ФБД не существует простого метода, чаще всего применяется аппарат Фурье. Также ФБД удобно определять с помощью параметра Хёрста H , где $H \in (0; 1)$, который указывает на степень изрезанности графика. При $H = 0, 5$ ФБД совпадает с классическим броуновским движением.

Гауссовский процесс $X(t)$ называется ФНД с параметром H , если: $X(0) = 0$ и траектории $x(t)$ почти всегда непрерывны, и $\Delta X = X(t_2) - X(t_1)$ имеет гауссовское распределение с нулевым математическим ожиданием и дисперсией $\delta^2 \cdot (t_2 - t_1)^{2H}$.

Разработанный метод использует оценку ряда статистических параметров, основан на полу-контролируемой технике выявления аномалий и реализован в виде одного из модулей портативного программного обеспечения.

Данный алгоритм содержит следующие этапы: запуск программы; перехват трафика; обработка трафика разработанным алгоритмом; реакция на аномалию.

Рассмотрим алгоритм анализа трафика на аномальность. Он состоит из двух частей: расчёт эталонных значений (обучение) и обнаружение аномалий.

Расчёт эталонов строится на значении параметра Хёрста для некоторых показателей из заголовков пакетов. В качестве таких показателей будем принимать количественное значение данных типа TCP-SYN и UDP, ICMP, полученных за 1 секунду. Производится сбор нормального трафика в течение 240 секунд. Определяем количество пакетов определённого типа за каждую секунду. После чего производим разбиение всего временного промежутка четырьмя способами: по 5, 15, 60 и 120 секунд на 48, 16, 4 и 2 интервалов, соответственно. Для каждого способа разбиения в каждом полученном интервале производим расчёт размаха, среднего квадратического отклонения (СКО) выборки и значений параметра Херста, для расчёта которого был выбран RS-анализ.

После вычисления значений параметра Хёрста для каждого интервала, определяем их среднее значение и СКО σ . Поскольку почти 96% случайных величин попадают в промежуток $(-2S; 2S)$, поэтому будем проверять попадание параметров Хёрста в промежуток «два сигма» (вместо правила «трёх сигм»). В данном случае, при большом числе наблюдений, когда среднее значение всей выборки неизвестно, правило преобразуется в правило «двух S». Таким образом, значения трафика, не попадающие в него, могут быть признаны аномальными. Но поскольку $H > 0$, то рассматривается промежуток $(0; 2S)$. В случае если значение параметра для анализируемого трафика выйдет за границы диапазона, система предупредит о возможной аномалии, и дополнительно будет произведена проверка по расширенному диапазону $(-3S; 3S)$. Если и в этот промежуток параметр анализируемого трафика не попадёт, система сообщит об наличии аномалии. Необходимость расчёта параметров для четыре различных интервалов обусловлена тем, что в случае малого числа пакетов, которые могут быть получены на небольших временных отрезках, возможна ложная реакция, но, однако, не стоит игнорировать данные на этих числовых промежутках, так как при появлении аномалии скорость её обнаружения является критичной.

Для оценки работоспособности метода было разработано консольное приложение на языке Python на ОС Kali Linux. Данные для расчёта эталонов и непосредственного анализа собираются при помощи Tshark.

Тестирование. В первую очередь будет произведено обучение. Для генерации нормального трафика произведем имитацию нормальной для стандартного пользователя деятельности: запуск видео на видеохостинге, вход в почтовый ящик и некоторое количество поисковых запросов. После завершения расчёта эталонов, начинается этап обнаружения аномалий. Для этого продолжим ту же деятельность, которая производилась на этапе обнаружения на протяжении 4-х минут. На протяжении данного времени сообщений об аномалии выдано не было, так как данный трафик система после обучения считает нормальной.

Тестирование срабатываний на реальные аномалии. Самый простой способ такой проверки — симуляция атаки SYN-flood при помощи генератора пакетов hping3.

Генератор пакетов запущен в режиме flood, то есть пакеты генерируются с такой скоростью, с которой это только возможно. К сожалению, отправка пакетов не отслеживается в терминале, так как их слишком большое количество. После запуска данного генератора система сообщит об аномалиях.

Результаты проеденного тестирования позволяют сделать вывод, что метод работает исправно и может быть использован для обнаружения аномалий. Для повышения его эффективности рекомендуется проводить расчёты по нескольким флагам одновременно, а также использовать в комплексе с уже существующими методами.

3. Заключение

Разработанный статистический метод выявления аномалий сетевого трафика готов к практическому использованию и может применяться в качестве модуля как в системах обнаружения вторжений, так и в DLP-системах. Данный метод учитывает современные тенденции по построению математических моделей сетевого трафика — его фрактальность.

Литература

1. Унтеров Д. С. Разработка метода обнаружения аномалий сетевого трафика на границе ЛВС предприятия: маг. дис. / Д. С. Унтеров. — Л., 2016. — 80 с.
2. Mohiuddin A., Abdun Naser M., Jiankun H. A survey of network anomaly detection techniques // J. Network and Comp. App. — 2015. — No 60. — P. 21.
3. Monowar H. Network Anomaly Detection: Methods, Systems and Tools // IEEE Communications surveys & tutorials. — 2013. — Vol. 60(1). — Pp. 303–336.
4. Aisaa N., Mohamed G. Semi-Supervised Statistical Approach for Network Anomaly Detection // Proc. Comp. Sci. — 2016. — Vol. 83. — Pp. 1090–1095.

UDC 51-74

A Construction of a Method for Detecting Anomalies of Network Traffic

E. V. Karachanskaya*[†], N. I. Sosodova[†]

* *Department of Information Technologies and Systems
Far-East State Transport University
47 Serysheva str., Khabarovsk, 680027, Russian Federation*

[†] *Department of Mathematics
Pacific National University
136 Tikho-oceanskaya str., Khabarovsk, 680035, Russian Federation*

Email: elena_chal@mail.ru, sosodova.nadezhda@yandex.ru

The authors of the scientific work offer a new method for detecting anomalies of network traffic modeled by fractal Brownian motion. The method refers to a group of statistical methods and is based on a semi-controlled method for detecting anomalies. The algorithm for analyzing traffic for anomalies consists of two parts: the calculation of reference values (training) and the detection of anomalies. The calculation of the standards is based on the value of the Hurst parameter for some indicators from the packet headers. Several time intervals are used for observation to reduce the number of false positives. The method's work is reduced to the comparison of the indicators at the stage of calculation of standards, with indicators calculated in the process of monitoring the anomalies. This method can be used to monitor both anomalies in incoming and outgoing traffic, which allows it to be included both in the intrusion detection suite and in the DLP system.

Key words and phrases: Network anomaly, Fractal Brownian motion, Hurst exponents, RSAnalysis.

УДК 004.94

Имитационное моделирование мобильности участников соединения в гетерогенной сети беспроводной передачи данных с ограничением на пересечение границ и ненадежными объектами

В. В. Кименчежи*, Д. В. Козырев*[†]

* *Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклухо-Маклая, д. 6, Москва, Россия, 117198*

[†] *Институт проблем управления им. В.А. Трапезникова РАН
ул. Профсоюзная, д. 65, Москва, Россия, 117997*

Email: rikk_a@bk.ru, kozyrev_dv@rudn.university

Рассматривается модель перемещения и связи рабочих, беспилотников и статичных машин на складе, с учётом ограничений на пересечение границ области и ненадёжности (возможности разрядки) беспилотников. Между объектами возможны три типа соединений: прямое соединение (Device-to-Device - D2D), соединение через Wi-Fi роутер на борту беспилотников и соединение через базовую станцию (БС). Проведено имитационное моделирование системы, в результате которого получены графики, отражающие положение объектов внутри склада, изменение площади покрытия, динамику работы беспилотников и доля участия рассматриваемых типов соединений в обеспечении связности движущихся участников соединения.

Ключевые слова: дискретно-событийное моделирование, надёжный Интернет вещей (RIoT), беспроводная передача данных, гетерогенная сеть.

1. Введение

Беспилотное управление и роботизация производства являются на данный момент наиболее востребованными и экономически интересными проектами в области высоких технологий для бизнеса [1]. Развитие технологии Интернета вещей даёт возможность повышения стабильности для общества и производства. Применение такого рода технологий позволяет сократить расходы ресурсов и оптимизировать работу склада. Одновременно с новыми возможностями использование таких технологий порождает и ряд требований [2], основное из которых — обеспечение стабильного надежного соединения между объектами Интернета вещей с малыми задержками передачи и высокой доступностью [3, 4].

Целью работы является разработка и программная реализация имитационной модели [5] мобильности агентов связи на складе и их взаимодействия посредством разных технологий беспроводной передачи данных с учётом ограничений на пересечение участниками соединения границ области и возможности разрядки беспилотников.

2. Описание задачи и исследуемой системы

В качестве объекта имитационного моделирования индустриальной автоматизации рассматривается склад площадью [100, 100] метров. Внутри склада располагаются 5 беспилотных аппаратов, 100 статичных машин и 30 мобильных сотрудников склада. Беспилотники и люди двигаются согласно модели Random Walk (RW) с малыми скоростями и в случайных направлениях. При этом они имеют ограниченный заряд аккумулятора и «отталкиваются» от границ области. Также имеется БС, обеспечивающая покрытие семидесяти процентов территории склада.

3. Результаты имитационного моделирования

Для реализации поставленной выше задачи была разработана программа на языке Python. На начальном этапе пользователю предоставляется возможность ввести значение модельного времени (т.е. количество итераций цикла), шаг (расстояние, на которое может сместиться объект), радиус покрытия беспилотника, радиус покрытия D2D. В результате работы программы рассчитывается общее количество объектов, имеющих доступ к беспилотникам, а также отдельно количество людей и количество машин, которым доступны различные типы связи на каждом шаге. Для нахождения площади покрытия общая площадь склада делится на равные квадраты, для каждого из которых находится его центр, затем подсчитывается количество центров квадратов, вошедших в зону покрытия беспилотниками. Полученная величина, умноженная на площадь квадрата, и будет составлять площадь общего покрытия беспилотниками и БС. Рис. 1 отображает пример состояния модели. Синие точки — машины, красные — беспилотники, зелёные — люди. Зелёные окружности — зоны покрытия связью активными беспилотниками, чёрная окружность — зона покрытия БС.

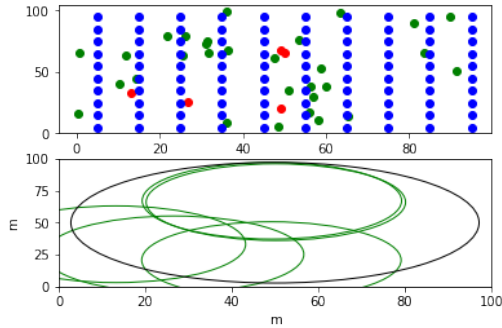


Рис. 1. Пример состояния имитационной модели

Программа также позволяет построить графики, отображающие соотношение разных типов связи по количеству использующих их агентов в процентах (рис. 2), количество беспилотников в работе (рис. 3), процент площади покрытия (рис. 4).

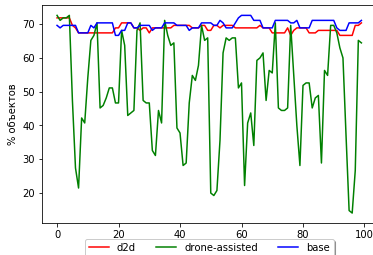


Рис. 2. Соотношение типов соединения

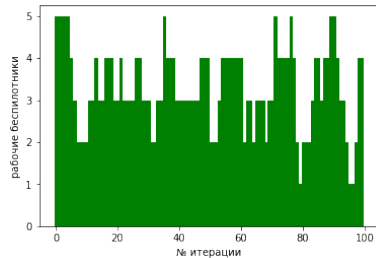


Рис. 3. Траектории процесса, описывающего работу беспилотников

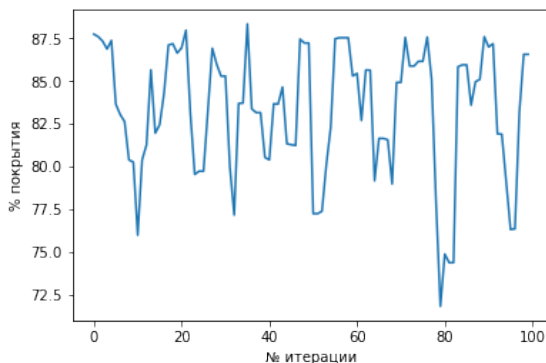


Рис. 4. Площадь зоны покрытия беспроводной связью беспилотниками и БС

4. Выводы

В результате работы программы были получены графики, позволяющие оценить площадь зоны покрытия связью рассматриваемых агентов. Сопоставление рисунков 2 – 4 позволяет оценить зависимость площади зоны покрытия беспроводной связью от количества беспилотников, находящихся в активном состоянии. Программа позволяет также визуализировать состояние системы в каждый момент времени, а также оценить прибавочную площадь покрытия беспроводной связью и рассчитать средние показатели по каждому типу соединения.

Литература

1. Ericsson, “Cellular Networks for Massive IoT”, White paper, 2015.
2. Кименчежи В.В., Козырев Д.В. Имитационное моделирование мобильности участников соединения в гетерогенной сети беспроводной передачи данных // Молодежная научная школа по прикладной теории вероятностей и телекоммуникационным технологиям (АРТСТ–2017) = 2nd International School on Applied Probability Theory Communications Technologies (ARTCT–2017): материалы молодежной научной школы. Россия, Москва, 23–27 октября 2017 г.; под общ. ред. К. Е. Самуйлова, Е. А. Кучерявого, А. Н. Дудина. — Москва: РУДН, 2017. — С. 138–141.
3. ITU-R M.2083-0, “IMT Vision — Framework and overall objectives of the future development of IMT for 2020 and beyond”, tech. rep., September 2015.
4. Kozyrev D., Kimenchezhi V., Houankpo H.G.K. Reliability Calculation of a Redundant Heterogeneous System with General Repair Time Distribution // Прикладные проблемы в теории вероятностей и математической статистике в области телекоммуникаций = Applied problems in theory of probabilities and mathematical statistics into telecommunications. Труды XI Международного семинара / под ред. Д. Аранити, К. Е. Самуйлова, С. Я. Шоргина. — М: РУДН, 2017. — С. 12.
5. Лоу А., Кельтон В. Имитационное моделирование. Классика ЦС, БХВ, 2004, 3 изд.

UDC 004.94

Mobility Simulation of Connected Objects in a Heterogeneous Wireless Data Transmission Network

V. V. Kimenchezhi*, D. V. Kozyrev*[†]

* *Department of Applied Probability and Informatics
Peoples' Friendship University of Russia*

6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation

[†] *V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences
Profsoyuznaya str. 65, Moscow, Russian Federation*

Email: rikk_a@bk.ru, kozyrev_dv@rudn.university

We consider a simulation model of a storage with three types of wireless connections: direct (D2D), drone-assisted and base. In this case, three types of mobility agents are considered: workers, drones and stationary machines. As a result of research we propose a software which generates graphical representation of mobility of objects inside the storage with given parameters, to assess the network coverage area and allows to study the IoT connectivity issues.

Key words and phrases: mathematical modeling and simulation, reliable Internet of Things (RIoT), wireless communication, heterogeneous system.

УДК 004.4

Гармонизация параметров IMS методом взаимной оптимизации по критериям производительности и потерь

С. Ш. Кутбитдинов*, В. В. Лохмотко[†], С. Р. Рудинская[‡]

* ГУП Центр научно-технических и маркетинговых исследований «UNICON.UZ»,
Узбекистан

[†] Санкт-Петербургский государственный университет телекоммуникаций
им. проф. М. А. Бонч-Бруевича, Россия

[‡] Белорусская государственная академия связи, Республика Беларусь

Email: kutbitdinov@unicon.uz, lvv1950@tut.by, sabina.rudin@mail.ru

Предлагается метод гармонизации (согласованной оптимизации) параметров IP Multimedia Subsystem, основанный на совместном решении задачи максимизации производительности и ограничительной задачи по качеству обслуживания.

Ключевые слова: гармонизация, модель Норроста, параметр Херста, коэффициент подобия, самоподобие, взаимные задачи, связанные переменные, IP Multimedia Subsystem.

1. Введение

Под гармонизованной сетью понимается сеть, в которой обеспечено синергетическое согласование (отсутствие диспропорций) топологических, потоковых и маршрутных показателей при условии достижения заданных целевых требований и выполнения проектных ограничений.

В докладе рассматривается задача согласования внешних параметров (нормативов по качеству обслуживания) с внутренними параметрами (пропускной способностью, объемами своевременно обслуженного сигнального трафика, загрузкой серверов и т.п.) сети нового поколения NGN/IMS (IP Multimedia Subsystem).

В операционном плане гармонизация представляет собой совместную оптимизацию параметров IP подсистемы по нескольким критериям, отличающуюся от классической многокритериальной оптимизации условием совпадения локальных оптимумов.

2. Постановка задачи

В качестве модели сети выбрана π -фазная СеМО $fB/M/1/\infty$ [1, 2] с интенсивностью μ обслуживания сервера и интенсивностью λ входящего самоподобного трафика типа «фрактальное броуновское движение».

Предполагаются заданными: IP Multimedia Subsystem с числом серверов π ; норма P_z на вероятность потери пакета; нормативное время t_z по доставке пакета; коэффициент подобия m [2], показывающий, во сколько раз задержке самоподобного трафика допускается превысить аналогичную задержку пуассоновского трафика. Требуется найти условия достижения предельной производительности G_{\max} (в числе своевременно доставленных пакетов) IP подсистемы и согласованную с ней загрузку ρ^* серверов (в том числе интенсивности λ и μ) при условии выполнения ограничений по качеству обслуживания (P_z и t_z).

3. Методология взаимной оптимизации

Основывается на идеях теории взаимности, сформулированной и доказанной А. Г. Аганбегяном и К. А. Багриновским, и сводится к совместной оптимизации ресурсных и потоковых параметров IP подсистемы по критериям вероятности P потери пакета и производительности G сети. Выделяются две взаимные задачи:

1. Ограничительная задача оптимизации параметров λ и μ под заданную норму P_z потерь с целевой функцией-лимиттером

$$P(\pi, \lambda, \mu) = P_z.$$

2. Экстремальная задача оптимизации параметра λ , доставляющая максимум G_{\max} производительности при целевой функции

$$G(\pi, \lambda, \mu) = \lambda(1 - P) \rightarrow \max. \quad (1)$$

Взаимным задачам свойственно:

- наличие общей части множества Ω допустимых решений. В данном случае функция $P(\pi, \lambda, \mu)$ является компонентой экстремальной функции $G(\pi, \lambda, \mu)$ производительности;
- совпадение (которое может быть далеко не всегда) решений взаимных задач по параметрам λ и μ ;
- различие природы критериев оптимизации. Если в 1-й задаче критерий является ограничивающим и безразмерным, то в 2-й экстремальным и измеряемым («пакет/с»).

4. Гармонизация параметров IMS

Для формализации вероятности P потерь в практике сетевого моделирования принято использовать дополнительную функцию распределения (д.ф.р.) случайной величины. Путем морфологического анализа в [2] показано:

- в среде двойного пуассоновского процесса система массового обслуживания $fB/M/1/\infty$ с самоподобным входящим трафиком сводится к экспоненциальной СМО $M/M/1/\infty$ с пуассоновским входящим потоком;
- средняя задержка W_n доступа (время ожидания) СМО $fB/M/1/\infty$ представляется геометрической прогрессией вида $W_n = W_1 \cdot q^{n-1}$ с дробным индексом n ;
- первый член $W_1 = \rho/(1 - \rho)/\mu$ прогрессии соответствует среднему времени ожидания пакета в очереди экспоненциальной СМО $M/M/1/\infty$;
- характеризующий темп роста прогрессии знаменатель q соответствует дисперсии $D = \rho \cdot (1 - \rho)^{-2}$ числа пакетов в СМО $M/M/1/\infty$;
- индексная переменная $n = 0, 5/(1 - H)$ является функцией параметра H Хёрста [1];
- коэффициент m [2] подобия

$$m = q^{n-1} = D^{n-1} = [\rho \cdot (1 - \rho)^{-2}]^{n-1}, \quad n = 0, 5/(1 - H), \quad 0, 5 \leq H < 1, \quad (2)$$

является коэффициентом преумножения (мультипликатором) средней длины «экспоненциальной» очереди до размера очереди СМО $fB/M/1/\infty$.

На базе д.ф.р. классического экспоненциального распределения с учетом (2) при предположении $\mu^{-1} \ll W_n$ строится аппроксимация функции потерь

$$P \approx \pi \cdot \exp [-(\mu - \lambda) \cdot t/m] = P_z. \quad (3)$$

Применительно к потоковым задачам оптимальное значение трафика λ^* определяется совместным решением (1)–(3) при условии $P = P_z$ и $t = t_z$ при заданных значениях m, P_z, t_z, μ . В ресурсных задачах определению подлежит оптимальная интенсивность μ^* обслуживания при известных параметрах m, P_z, t_z, λ .

При $m = 1$ (для пуассоновского входящего потока) из (3) следует уравнение связи

$$\mu - \lambda = (\ln \pi - \ln P)/t. \quad (4)$$

Совместное решение (4), связывающего остаточную пропускную способность $(\mu - \lambda)$ сервера с нормативными параметрами, и уравнения $\partial G / \partial \lambda = 0$, соответствующего условию достижения максимума производительности (при фиксированном μ), дает гармонизированные значения интенсивностей входящего потока λ^* и обслуживания μ^* ,

$$\lambda^* = \frac{1 - P_z}{t_z P_z}, \quad \mu^* = \frac{1}{t_z} \left(\frac{1 - P_z}{P_z} + \ln \pi - \ln P_z \right), \quad (5)$$

а также максимальную производительность G_{\max} , как функцию нормативов P_z и t_z ,

$$G_{\max} = (1 - P_z)^2 \cdot t_z^{-1} \cdot P_z^{-1}.$$

Значения других показателей (средней длины очереди, средней задержки пакета, числа пакетов в сети, загрузки сервера и т.д.) легко рассчитываются на базе (5).

5. Основные результаты

Поскольку максимум производительности IMS достигается при неоптимальных QoS-показателях была сформулирована задача их гармонизации (согласованной оптимизации со связанными переменными μ и λ). Показано, что производительность IMS представляется линейной функцией (дающей семейство решений), зависящей от интенсивности λ входящего трафика и остаточной пропускной способности $(\mu - \lambda)$ сервера. Найденные условия согласования позволяют легко пересчитывать внешние (инфраструктурные) параметры (в данном случае QoS-нормативы) в сетевые показатели (загрузку, задержку, производительность и т.д.).

Литература

1. Norros Ilkka. A storage model with self-similar input // *Queueing Systems*. — 1994. — Vol. 16, no. 3. — Pp. 387–396.
2. Кутбитдинов С. Ш. Рекуррентный пуассоновский аналог Норросовской модели СМО с самоподобным входящим трафиком / С. Ш. Кутбитдинов, В. В. Лохмотко, С. Р. Рудинская // *Инфокоммуникации. Сети-Технологии-Решения*. — 2017. — № 3. — С. 10–16.

UDC 004.4

Harmonization of IMS Parameters by Mutual Optimization by Criteria for Performance and Losses

S. Sh. Kutbitdinov*, V. V. Lokhmotko†, S. R. Rudinskaya‡

* *GUP Center for Scientific and Technical and Marketing Research "UNICON.UZ", Uzbekistan*

† *St. Petersburg State University of Telecommunications. prof. M.A. Bonch-Bruevicha*

‡ *Belarusian State Academy of Communications, Republic of Belarus*

Email: kutbitdinov@unicon.uz, lvv1950@tut.by, sabina.rudin@mail.ru

The method of harmonization (coordinated optimization) of IP Multimedia Subsystem parameters is proposed, based on the joint solution of the task of maximizing performance and restricting the problem of quality of service.

Key words and phrases: harmonization, Norros model, Hurst parameter, similarity coefficient, self-similarity, mutual tasks, related variables, IP Multimedia Subsystem.

УДК 621.39

Анализ установления соединения пар передатчик-приемник на mmWave

Е. Д. Макеева*, П. А. Харин*, Н. А. Поляков*, Е. В. Маркова*,
И. А. Гудкова*[†], О. С. Галинина*

* Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198

[†] Федеральный исследовательский центр «Информатика и управление» РАН
ул. Вавилова, д. 44, к. 2, Москва, Россия, 119333

Email: len16730637@yandex.ru, gruzavjeg@mail.ru, goto97@mail.ru,
markova_ev@pfur.ru, igudkova@sci.pfu.edu.ru, olga.galinina@gmail.com

В связи с экспоненциальным ростом трафика, передаваемого в беспроводных сетях, возникает проблема нехватки частотного диапазона, приводящая к необходимости разработки новых технологий, требуемых для качественной передачи данных. Одной из таких технологий является система mmWave, использующая миллиметровый диапазон частот, позволяющий передавать данные на высоких скоростях. Особенностью данной системы является низкая дальность сигнала, а также достаточно точное определение диапазона, скорости и угла его распространения. В работе рассмотрена задача определения области действия сигнала пары передатчик – приемник на плоскости.

Ключевые слова: mmWave, границы распространения сигнала.

1. Введение

В данной работе рассматривается произвольная область на плоскости, в которой равномерно и непрерывно возникают пары точек. Обозначим эти точки A и B . Пусть точка A передает данные (передатчик), точка B принимает (приемник). Сигналы между точками распространяются в пределах некоторой области. Задача исследования состоит в том, чтобы определить максимальное количество пар передатчик-приемник, которые могут возникнуть в заданной плоскости так, чтобы области распространения сигнала каждой пары не пересекались друг с другом, в противном случае передача сигнала от вновь возникнувшей пары будет заблокирована. Для этой цели в работе предложен алгоритм определения границ распространения сигнала между произвольной парой передатчик-приемник.

2. Анализ границ распространения сигнала

Пусть точка A имеет координаты (x_1, y_1) , точка $B - (x_2, y_2)$. Угол распространения сигнала равен α , дальности действия передатчика и приемника одинаковы — R (рис. 1). Обозначим расстояние между точками A и B как d . Будем считать, что в случае, когда расстояние d больше, чем дальность действия R , то соединение не может быть установлено. Таким образом, для установления соединения необходимо выполнение неравенства (1)

$$\sqrt{(x_1 - x_2)^2 - (y_1 - y_2)^2} < R. \quad (1)$$

Перейдем к определению границ распространения сигнала между точками A и B , согласно рис. 1 — определению границ треугольника AEC и BDF .

В качестве примера рассмотрим анализ только одного из них — AEC , для этого разобьем его на два треугольника ABE и ABC .

Шаг 1. Найдем длину $AB=d$ по формуле (2)

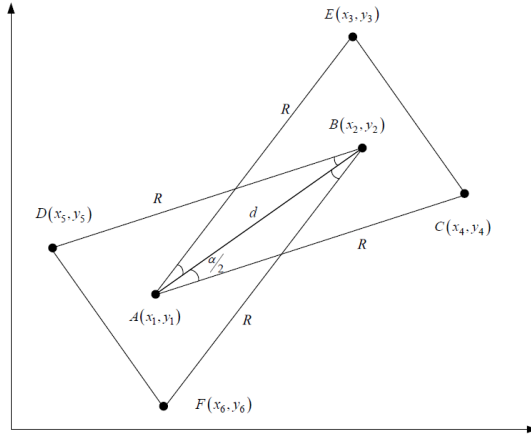


Рис. 1. Построение области распространения сигнала между точками A и B

$$d = \sqrt{(x_1 - x_2)^2 - (y_1 - y_2)^2}. \quad (2)$$

Шаг 2. Рассчитаем длины сторон BE и BC по теореме косинусов (3)

$$BE = BC = \sqrt{R^2 + d^2 - 2Rd \cos \frac{\alpha}{2}}. \quad (3)$$

Шаг 3. Составим систему уравнений (4), решение которой позволяет определить значения координат точек E и C , т.е. пары (x_3, y_3) , (x_4, y_4)

$$\begin{cases} BE^2 = (x_2 - x_i)^2 - (y_2 - y_i)^2, \\ R^2 = (x_1 - x_i)^2 - (y_1 - y_i)^2, \end{cases} \quad (4)$$

где $i = 3, 4$.

Шаг 4. Запишем уравнения прямых AE , AC , EC , используя координаты точек A , E , C . Подставим координаты (x_1, y_1) , (x_3, y_3) , (x_4, y_4) в формулу уравнения прямой, проходящей через две точки (5)

$$\frac{y - y_j}{y_i - y_j} = \frac{x - x_j}{x_i - x_j}, \quad (5)$$

где $i = 3, 4$, $j = 1, 3$.

Шаг 5. Выпишем неравенства (6), определяющие область треугольника AEC

$$\begin{cases} AE : y \leq \frac{(x-x_1)(y_3-y_1)}{x_3-x_1} + y_1; \\ AC : y \geq \frac{(x-x_1)(y_4-y_1)}{x_4-x_1} + y_1; \\ CE : y \leq \frac{(x-x_3)(y_4-y_3)}{x_4-x_3} + y_3. \end{cases} \quad (6)$$

Расчет границ второго треугольника BDF проводится по аналогии.

3. Заключение

В работе рассмотрен пример возникновения пар передатчик-приемник в некоторой области на плоскости. Определены границы распространения сигнала между произвольной парой передатчик-приемник. Полученные результаты позволят определить максимальное количество пар передатчик-приемник, которые могут возникнуть в заданной плоскости, так чтобы области распространения сигнала каждой пары не пересекались друг с другом, т.е. так чтобы при возникновении новой пары не происходила блокировка сигнала.

Благодарности

Публикация подготовлена при поддержке программы РУДН «5-100» и при финансовой поддержке РФФИ в рамках научных проектов № 17-07-00142, 18-07-00567.

UDC 621.39

Analysis of the Connection Setup of Transmitter-Receiver Pairs on mmWave

E. D. Makeeva*, P. A. Harin*, N. A. Polyakov*, E. V. Markova*,
I. A. Gudkova*[†], O. S. Galinina*

* *Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

[†] *Institute of Informatics Problems, FRC CSC RAS
44-2 Vavilova Str., Moscow 119333, Russian Federation*

Email: len16730637@yandex.ru, gruzavjeg@mail.ru, goto97@mail.ru,
markova_ev@pfur.ru, igudkova@sci.pfu.edu.ru, olga.galinina@gmail.com

In connection with the exponential growth of traffic transmitted in wireless networks, there arises the problem of lack of a frequency range, which leads to the need to develop new technologies required for high-quality data transmission. One such technology is the mmWave system, using a millimeter frequency range, which allows data to be transmitted at high speeds. The peculiarity of this system is the low signal range, as well as a sufficiently accurate determination of the range, speed and angle of its propagation. The problem of determining the range of the signal of a transmitter-receiver pair on a plane is considered in this paper.

Key words and phrases: mmWave, signal propagation boundaries.

УДК 519.217

Имитационное моделирование уличных точек доступа, функционирующие на миллиметровом диапазоне частот

Е. А. Мачнев, В. О. Бегишев

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: egork1996@yandex.ru, begishev_vo@rudn.university

Технологии беспроводной связи пятого поколения, работающие в миллиметровом диапазоне, ставят перед разработчиками систем все новые трудности. Частое изменение мощности принимаемого сигнала, генерируемый различными мобильными устройствами в сети, приводит к потерям пакетов. Для решения этой проблемы 3GPP предложил множественное соединение, когда мобильное устройство одновременно поддерживает соединения с несколькими базовыми станциями. Помимо этого, мобильное устройство может двигаться в пространстве, что усложняет задачу, так как необходимо постоянное переключение. В этой работе мы предлагаем основу для симуляции одного из самых сложных сценариев — движение автомобиля по улице с несколькими полосами. В работе будут представлены такие метрические характеристики, как потери пакетов данных, в зависимости от плотности расположения базовых станций и степень многосвязности.

Ключевые слова: mmWave (millimeter-wave), миллиметровый диапазон, имитационное моделирование.

1. Введение

Технология беспроводного доступа, работающая в миллиметровом диапазоне (mmWave), призвана стать решающим шагом в эволюции существующих сетей связи. Процесс стандартизации в 3GPP почти завершен, однако существует проблема оптимизации таких сетей в различных сценариях использования. Вместе с уникальными возможностями mmWave приносят и трудности разработчикам систем. Так, например, для стабильной работы между мобильным устройством и базовой станцией должна быть прямая видимость, а так как большинство объектов в сети мобильны, то перекрыть сигнал может что угодно, поэтому поведение системы непредсказуемо. Для решения этой проблемы 3GPP предложил поддерживать одному мобильному устройству сразу несколько соединений с различными базовыми станциями. В каждый раз, когда мощность сигнала падает ниже определенного уровня, мобильное устройство переключается на одну из резервных точек доступа.

Преимущества многосвязности обсуждались в [1], где авторы рассматривали вероятность отказа при установлении соединений с точками доступа. Эти успехи были дополнительно оценены в [2], где было проведено моделирование системы улицы. В [3] рассмотрен частный случай, когда соединение осуществлено с двумя точками доступа. Распределение отношения сигнал-шум в сетях mmWave с возможностью мультиподключения было получено в [4]. Сравнение стратегий переключения точек доступа для сетей mmWave было выполнено в [5]. Все вышеупомянутые исследования рассматривают открытые пространства со статичными объектами, а также само мобильное устройство не рассматривалось как подвижное. Основной причиной таких исследований является то, что математический анализ динамического движения довольно сложен. Если учитывать и мобильность устройства и движение окружающих объектов, то расчеты становятся непростыми. В работе мы рассматриваем случай движущегося по определенной улице автомобиля, который получает обслуживание от точек доступа, установленных вдоль дороги, а также мы учитываем наличие блокираторов сигнала в виде автомобилей, движущихся как попутно нашему автомобилю, так и на встречу. Мы показываем, что при геометрическом подходе сценарий имитационного моделирования может быть сильно упрощен, что

позволит моделировать такие сети и рассчитывать для них показатели эффективности. Предлагаемый сценарий может быть использован для более сложных сценариев систем mmWave. В следующем разделе мы вводим системную модель рассматриваемого сценария улицы, но численный анализ для этой системы будем представлять в следующих работах. Выводы приводятся в последнем разделе.

2. Системная модель

Рассмотрим модель улицы, изображенной на Рис. 1. Предполагается, что улица имеет $N = 4$ полос, где N всегда четно. Ровно первая половина полос движется в одном направлении, а вторая половина — в противоположном. Все полосы имеют одинаковую ширину w . Покажем случай, когда наше транспортное средство движется по одной из внутренних полос улицы с постоянной скоростью V_U . Высота машины считается постоянной и равна h_U .

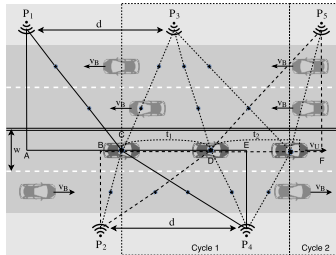


Рис. 1. Сценарий улицы с mmWave точками доступа

Предположим, что наше транспортное средство не блокирует само себя. Точки доступа расположены по обе стороны улицы на одинаковом расстоянии d , например, на фонарных столбах на одинаковой высоте h_A . Точки доступа по разные стороны улицы расположены таким образом, что они образуют равнобедренные треугольники. Прямая видимость между точкой доступа и нашим транспортным средством может быть блокирована а) большими статичными объектами, например зданиями, б) небольшими движущимися объектами, например людьми или машинами.

В нашем случае, мы рассматриваем только автомобили, так как точки доступа расположены по краям дороги и на линии сигнала к нашему автомобилю не может быть никакого здания. Блокаторы сигнала (автомобили) представлены одномерными прямоугольниками, расположенными вдоль оси каждой полосы и имеющие постоянную скорость V_B . Длина и высота автомобилей — случайные величины с функциями распределения $F_{L,B}(x)$ и $F_{H,B}(x)$. Центры сторон этих прямоугольников образуют на полосе пуассоновский процесс с интенсивностью λ_B . Предполагается, что перекрытие прямой видимости другим автомобилем ухудшает сигнал на 20 дБ. Сила потери сигнала представлена (1):

$$L(d) = \begin{cases} 32.4 + 21.0 \log(d) + 20 \log f_c, & \text{заблокировано} \\ 52.4 + 21.0 \log(d) + 20 \log f_c, & \text{незаблокировано,} \end{cases} \quad (1)$$

где d — расстояние от точки доступа до нашего транспортного средства, а f_c — частота полосы в ГГц. Для каждой точки доступа определены три зоны. Целевой машине разрешено использовать несколько одновременных соединений. Автомобиль поддерживает одновременную связь с M точками доступа, где M называется

степенью многосвязности. Время переключения в стандарте еще не описано, поэтому на данный момент мы предполагаем что оно происходит мгновенно. В описанной выше системе будем высчитывать среднее время, в течение которого точки доступа были заблокированы проезжающими автомобилями.

3. Заключение

В этой статье мы предложили основу для симуляции сценария улицы с использованием многосвязности 3GPP и наличием блокираторов на соседних полосах движения. Мы заметили, что связность целевого автомобиля повторяется циклически во время движения вдоль улицы, что значительно упрощает моделирование, оставляя только движение блокираторов, так как остальные события могут быть математически вычислены.

Литература

1. Michalopoulos D.S., Viering I, and Du L. User-plane multi-connectivity aspects in 5G, in 2016 23rd International Conference on Telecommunications (ICT), pp. 1–5, May 2016.
2. Tesema F. B., Awada A., Viering I., Simsek M., and Fettweis G. P. Mobility modeling and performance evaluation of multi-connectivity in 5G intra-frequency networks, in IEEE Globecom Workshops (GC Wkshps), pp. 1–6, December 2015.
3. Polese M., Giordani M., Mezzavilla M., Rangan S., and Zorzi M. Improved handover through dual connectivity in 5G mmWave mobile networks // IEEE Journal on Selected Areas in Communications, September 2017. — Vol. 35. — Pp. 2069–2084.
4. Arduengo A. J., Harlow R. L., Kline M. A stable crystalline carbene // J. Am. Chem. Soc. — 1991. — Vol. 113 (1). — Pp. 361–363. doi:10.1021/ja00001a054.
5. Ohmann D., Awada A., Viering I., Simsek M., and Fettweis G. P. Achieving high availability in wireless networks by inter-frequency multi-connectivity, in 2016 IEEE International Conference on Communications (ICC), pp. 1–7, May 2016.

UDC 519.217

Simulation Modeling of Street Access Points Operating on the Millimeter Frequency Band

Е. А. Мачнев, В. О. Бегисhev

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: egork1996@yandex.ru, begishev_vo@rudn.university

Technologies of 5G wireless communication operating in the millimeter range pose new challenges for system designers. Frequent changes received signal power generated by various mobile devices on the network result in packet loss. To solve this problem, 3GPP proposed a multiple connection when the mobile device simultaneously supports connections to several base stations. In addition, the mobile device can move in space, which complicates the task, since it is necessary to constantly switch. In this work, we offer the basis for simulating one of the most difficult scenarios - the movement of a car along a street with several lanes. The work will present such metric characteristics as the loss of data packets, depending on the density of the location of base stations and the degree of multiplicity.

Key words and phrases: mmWave (millimeter-wave), millimeter range, simulation.

УДК 519.2

Исследование наработки сенсорной сети полиграфического комбината до отказа

И. А. Мацкевич

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: ivan.matskevich.1994@gmail.com

Благодаря распространению концепции Интернета вещей и повышению доступности элементов, образующих сенсорные сети, мир промышленности стремительно превратился из аналогового в цифровой. Если раньше принцип работы устройства закладывался буквально в ходе создания и был абсолютно неизгладимым весь срок службы машины, то сегодня вся логика реализована программными методами и может быть изменена или настроена в соответствии с решаемыми задачами. В работе рассматривается настраиваемая сенсорная сеть, развернутая на полиграфическом комбинате, строится ее математическая модель и рассчитывается вероятность ее отказа.

Ключевые слова: интернет вещей, дерево отказов, бинарная диаграмма решений.

1. Введение

Объектом исследования данной работы является сенсорная сеть, которая развернута на полиграфическом комбинате. Целью системы является уменьшение нагрузки на инфраструктуру. Исследуемая система представляет из себя набор сенсоров, объединенных в узлы, которые в свою очередь собирают количественные или качественные данные и передают их на дальнейшую обработку.

Система состоит из трех подсистем:

- 1) система подсчета печатных пластин;
- 2) система мониторинга за котлом;
- 3) система контроля печатной машины.

Подсистемы 1 и 2 соответственно состоят из сенсоров, подключенных к контроллеру ESP8266 по Wi-fi. Подсистема 3 состоит из сенсоров, передающих данные по проводу на контроллер ATmega2560. Данные, собираемые с датчиков, обрабатываются программой, написанной на языке C++, и записываются в таблицу в облаке.

Целью исследования является расчет наработки данной системы до отказа на основе предложенной архитектуры и известного перечня приборов и технологий.

Для оценки времени наработки описанной системы до отказа был использован метод деревьев отказа, описанный в [1] и примененный для аналогичного расчета в [2]. По построенному дереву составляется бинарная диаграмма решений, позволяющая написать уравнение для оценки наработки описанной системы. При подстановке вероятностей отказов уравнение даёт вероятность работы системы (или ее наработку до отказа, узнав сколько система будет простаивать в заданный промежуток времени).

2. Математическая модель сенсорной сети

Мы рассматриваем систему в промышленной, «грязной» среде, полной различных помех. Поэтому при построении дерева отказов необходимо учитывать не только отказы компонентов системы (сенсоры и контроллеры), но и отказы, связанные с передачей данных с сенсоров на контроллеры и с контроллеров в таблицу в облаке. На рис. 1 изображен фрагмент дерева, показывающий отказы, связанные с типом соединения сенсора и контроллера.

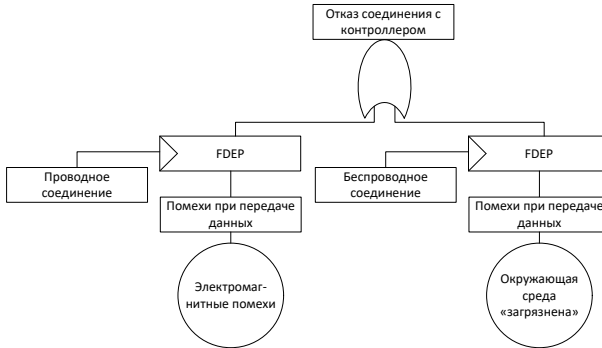


Рис. 1. Фрагмент дерева отказов

Основываясь на бинарной диаграмме решений, было составлено уравнение для получения вероятностей отказа системы (1):

$$\begin{aligned}
 P = & X_1 + (\bar{X}_1 \cdot X_2 \cdot X_3) + (\bar{X}_1 \cdot \bar{X}_2 \cdot X_4) + \\
 & + (\bar{X}_1 \cdot (X_2 \cdot \bar{X}_3 + \bar{X}_2 \cdot \bar{X}_4) \cdot X_5) + \\
 & + (\bar{X}_1 \cdot (X_2 \cdot \bar{X}_3 + \bar{X}_2 \cdot \bar{X}_4) \cdot \bar{X}_5 \cdot X_6) + \\
 & + (\bar{X}_1 \cdot (X_2 \cdot \bar{X}_3 + \bar{X}_2 \cdot \bar{X}_4) \cdot \bar{X}_5 \cdot \bar{X}_6 \cdot X_7). \quad (1)
 \end{aligned}$$

Ниже (см. табл. 1) представлена таблица с исходными данными, использующиеся для получения численного результата. Подставив вероятности, описанные в табл. 1, получим из (1):

$$\begin{aligned}
 P = & 0.00001 + (0.99 \cdot 0.66 \cdot 0.00001) + (0.99 \cdot 0.33 \cdot 0.00001) + \\
 & + (0.99 \cdot (0.66 \cdot 0.99 + 0.33 \cdot 0.99) \cdot 0.00001) + \\
 & + (0.99 \cdot (0.66 \cdot 0.99 + 0.33 \cdot 0.99) \cdot 0.99 \cdot 0.00001) + \\
 & + (0.99 \cdot (0.66 \cdot 0.99 + 0.33 \cdot 0.99) \cdot 0.99 \cdot 0.99 \cdot 0.00001) = 0.00004861985.
 \end{aligned}$$

В итоге получаем, что вероятность отказа рассмотренной системы равна 0.00004861985. После преобразований получаем, что сенсорная сеть на полиграфическом комбинате находится в состоянии отказа 25,5 минут в году.

3. Заключение

В результате описания сенсорной сети, анализа ее дерева отказов и бинарной диаграммы решений было получено уравнение, позволяющее оценить наработки данной системы до отказа. В дальнейшем можно раскрыть отказы, связанные не только с передачей данных с сенсоров на контроллеры, но и самих компонентов системы.

Таблица 1

Исходные данные для рассматриваемой системы

Обозначение	Параметр	Значение
X_1	вероятность отказа сенсора	0.00001
X_2	вероятность, что соединение между сенсором и контроллером беспроводное	0.66
X_3	вероятность ошибки при беспроводной передаче	0.00001
X_4	вероятность ошибки при передаче по проводу	0.00001
X_5	вероятность того, что контроллер недоступен	0.00001
X_6	вероятность отказа контроллера	0.00001
X_7	вероятность того, что облачный сервис оказался недоступен	0.02

Литература

1. Хенли Э. Дж., Кумamoto X. Надежность технических систем и оценка риска. — М.: Машиностроение, 1984.
2. Liudong Xing, Massarrah Tannous, Vinod M. Vokkarane, Honggang Wang and Jun Guo, Reliability Modeling of Mesh Storage Area Networks for Internet of Things // IEEE Internet of Things Journal, 2017.

UDC 519.2

Research on Working Time before Failure of the Sensor Network Installed on a Printing Plant

I. A. Matskevich

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: ivan.matskevich.1994@gmail.com

By spreading the concept of the Internet of things the world of industry has rapidly evolved from analog to digital. Before the principle of the device was laid literally during the creation and the entire lifetime of the machine was absolutely unshakable, today all logic is implemented by software methods and can be changed or tuned in accordance with the tasks to be solved. In the paper we consider a sensor network deployed at a printing plant, construct its mathematical model and calculate its failure probability.

Key words and phrases: internet of things, binary decision diagram, dynamic fault tree.

УДК 004.4

Численный анализ оптимальной скорости передачи данных в сети с технологией мультимедиа методом перебора

М. П. Половов*, В. А. Бесчастный*,
Д. Ю. Острикова*, И. А. Гудкова*[†]

* Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклухо-Маклая, д. 6, Москва, Россия, 117198

[†] Институт проблем информатики
Федеральный исследовательский центр «Информатика и управление» РАН
ул. Вавилова, д. 44, к. 2, Москва, Россия, 119333

Email: lojka10@list.ru, ostrikova_dyu, beschastnyy_va, gudkova_ia@rudn.university

В данной работе рассматриваются методы выделения подгрупп мультимедиа, позволяющие максимизировать совокупную, сбалансированную и уравновешенную скорости передачи данных в сетях LTE.

Ключевые слова: совокупная скорость передачи данных (ADR), сбалансированная скорость передачи данных (BDR), усредненная скорость передачи данных (FDR), задача оптимизации, выделение подгрупп мультимедиа.

1. Введение

Растущие объемы трафика мобильных приложений, использующих передачу потокового видео и файлов большого объема, заставляет задуматься о том, как можно оптимизировать процесс передачи данных от базовой станции до мобильного устройства. Одним из способов повысить скорость передачи данных является применение технологии мультимедиа, которая позволяет передавать некоторой подгруппе пользователей, заинтересованным в одном контенте, данные с использованием минимального количества ресурсов сети. В зависимости от технологий прикладного уровня и потребностей пользователей, может применяться один из трех методов максимизации соответствующей целевой характеристики: совокупная скорость передачи данных (англ. Aggregate Data Rate) [1], сбалансированная скорость передачи данных (англ. Balanced Data Rate) и усредненная скорость передачи данных (англ. Fair Data Rate). Алгоритмы, рассмотренные в данной статье, основаны на выделении подгрупп мультимедиа и распределении ресурсов сети между ними.

2. Постановка задачи

В сети LTE в зависимости от уровня качества принимаемого сигнала можно выделить M зон, каждой из которых сопоставлен один индекс CQI (англ. Channel Quality Index). В зависимости от значения CQI пользователь, находящийся в соответствующей зоне, получает различную скорость передачи данных из расчета на один выделенный ресурсный блок. В случае простейшей модели соты без препятствий на пути от базовой станции к пользователям CQI-зоны представлены в виде колец. Объединяя смежные кольца, выделяются подгруппы мультимедиа. Так, скорость на один ресурсный блок для всех пользователей в подгруппе рассчитывается как минимальная скорость передачи среди всех пользователей этой подгруппы.

3. Описание алгоритмов и численный анализ

Введем обозначения: n — массив, содержащий кол-во пользователей в каждой группе; b — массив, содержащий максимальную скорость на блоке в каждой группе; $maxR$ — максимальное количество колец; $maxB$ — максимальное кол-во ресурсных блоков.

<p>ADR($n, b, maxR, maxB$)</p> <p>– Выход: $maxADR$ – максимальная ADR.</p> <p>$maxADR \leftarrow -1$</p> <p>for всех групп u</p> <p> if $n[u] = 0$</p> <p> then удалить $n[u], b[u]$</p> <p>for всех разбиений колец на группы r</p>	<p>for всех разбиений ресурсных блоков на группы k</p> <p> $ADR \leftarrow 0$</p> <p> for всех групп u</p> <p> $ADR \leftarrow ADR + k[u] * n[u] * b[u]$</p> <p> if $ADR > maxADR$</p> <p> then $maxADR \leftarrow ADR$</p>
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Рис. 1. ADR метод

<p>BDR($n, b, maxR, maxB$)</p> <p>– Выход: $maxBDR$ – максимальное наименьшее перемножение k и b для каждой подгруппы</p> <p>$maxBDR \leftarrow -1$</p> <p>for всех групп u</p> <p> if $n[u] = 0$</p> <p> then удалить $n[u]$ удалить $b[u]$</p>	<p>for всех разбиений колец на группы r</p> <p> for всех разбиений ресурсных блоков на группы k</p> <p> if $maxBDR = -1$</p> <p> then $maxBDR \leftarrow k[0] * b[0]$</p> <p> for всех групп u</p> <p> $FDR \leftarrow k[u] * b[u]$</p> <p> if $FDR < maxBDR$</p> <p> then $maxBDR \leftarrow FDR$</p>
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Рис. 2. BDR метод

<p>FDR($n, b, maxR, maxB$)</p> <p>– Выход: $minDifFDR$ – разница между максимальной и минимальной скоростями</p> <p>$minFDR \leftarrow -1$</p> <p>$maxFDR \leftarrow -1$</p> <p>$minDifFDR \leftarrow -1$</p> <p>for всех групп u</p> <p> if $n[u] = 0$</p> <p> then удалить $n[u]$ удалить $b[u]$</p> <p>for всех разбиений колец на группы r</p>	<p>for всех разбиений ресурсных блоков на группы k</p> <p> if $minFDR = -1$</p> <p> then $minFDR \leftarrow k[0] * b[0]$</p> <p> for всех групп u</p> <p> $FDR \leftarrow k[u] * b[u]$</p> <p> if $FDR < minFDR$</p> <p> then $minFDR \leftarrow FDR$</p> <p> if $FDR > maxFDR$</p> <p> then $maxFDR \leftarrow FDR$</p> <p> if $minDifFDR = -1$ или $maxFDR - minFDR < minDifFDR$</p> <p> then</p> <p> $minDifFDR \leftarrow maxFDR - minFDR$</p>
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Рис. 3. FDR метод

На рис. 1–3 алгоритмы представлены в виде псевдокода. Для предложенных алгоритмов в качестве примера расчета в работе представлены результаты их работы в табл. 1.

4. Выводы

Алгоритм максимизации ADR принципиально отличается от BDR и FDR тем, что он зависит от количества пользователей (то есть от n), чего нельзя сказать про два других алгоритма. Максимизация ADR позволяет повысить пропускную способность сети и может применяться в тех случаях, когда нужно максимально

быстро передать большой объем данных пользователям сети. BDR и FDR могут применяться там, где установлены строгие нормы качества обслуживания.

Пример расчета для системы с пятью уровнями

Таблица 1

Входные параметры	Исходные данные	ADR	BDR	FDR
n	2,3,0,1,4			
b	100,80,60,40,20	2000	20	40
$maxR$	5			
$maxB$	10			
n	5,3,10			
b	100,80,60,40,20	5240	60	60
$maxR$	10			
$maxB$	10			

Благодарности

Публикация подготовлена при поддержке Программы РУДН «5-100» и при финансовой поддержке РФФИ в рамках научных проектов № 17-07-00845 и № 18-07-00576.

Литература

- Ostrikova D., Rinaldi F., Beschastnyi V., Gudkova I., Militano L., Araniti G., Iera A., Samouylov K. Analytical Model for Multicast Subgrouping in 5G-Mobile eMBMS Environment // Congress Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 2017 9th International, 6–8 Nov. 2017. <https://doi.org/10.1109/ICUMT.2017.8255188>

UDC 004.4

Numerical Analysis of the Optimum Speed of Data Transmission in the Network with Multiplier Technology by Bruteforce

M. P. Polovov*, V. A. Beschastnyi*, D. Yu. Ostrikova*, I. A. Gudkova*†

* Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
Miklukho-Maklaya str. 6, Moscow, 117198, Russian Federation

† Institute of Informatics Problems, FRC CSC RAS
44-2 Vavilov st., Moscow, 119333, Russian Federation

Email: lojka10@list.ru, ostrikova_dyu, beschastnyy_va, gudkova_ia@rudn.university

In this work we consider methods for multicast subgrouping to maximize the aggregate, balanced, and fair data rates in LTE networks.

Key words and phrases: Aggregate Data Rate (ADR), Balanced Data Rate (BDR), Fair Data Rate (FDR), optimization problem, multicasting subgrouping.

УДК 004.4

Математическая модель среднего времени передачи данных подвижным пользователям в сети LTE

В. А. Разгоняев*, Е. В. Мокров*, К. Е. Самуйлов*†

* Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198

† Федеральный исследовательский центр «Информатика и управление» РАН
ул. Вавилова, д. 44, к. 2, Москва, Россия, 119333

Email: vladislav.razgonyayev@gmail.com, mokrov_ev@rudn.university, samuylov_ke@rudn.university

В данной работе рассмотрена математическая модель качества передачи данных пользователями, которые перемещаются в пределах соты мобильной связи. Сота разделена на зоны, которые имеют разное качество обслуживания (channel quality indicator, CQI). Чем дальше CQI от центра соты, тем меньше скорость передачи данных. Целью работы является анализ среднего времени в зависимости от объема передаваемых данных подвижным пользователям.

Ключевые слова: мультимедиа; гетерогенная беспроводная сеть; показатели качества обслуживания; управляющий марковский процесс; среднее время передачи; двойственный процесс; сеть массового обслуживания.

1. Введение

Рассмотрим соту подвижной связи с N абонентами и M уровнями качества обслуживания, зависящими от местоположения абонентов в соте. Сота разделена на зоны, которые имеют разное качество обслуживания (channel quality indicator, CQI). Пользователи движутся согласно движению Леви и могут переходить из одной зоны обслуживания в другую.

2. Основная часть

Введем управляющий марковский процесс $\mathbf{X}(t) = (X_1(t), X_2(t), \dots, X_M(t))$. Такой управляющий процесс $\mathbf{X}(t)$ имеет пространство состояний:

$$X = (n_1, \dots, n_M) \mid n_i = 0, \dots, N, \quad i = 1, \dots, M, \dots \sum_{i=1}^M n_i = N.$$

При этом множество состояний процесса имеет мощность $|X| = C_{M+N-1}^N$. Процесс $\mathbf{X}(t)$ является марковским. Его можно рассматривать как процесс, описывающий некую замкнутую сеть массового обслуживания с N заявками и M узлами неограниченной ёмкости. В этой сети длительности обслуживания заявок в узле i имеют экспоненциальное распределение с параметром τ_i^{-1} [1]. На рис. 1 представлен граф вероятностей переходов для одного пользователя в сети.

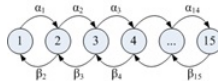


Рис. 1. Граф вероятностей переходов для одного пользователя в сети

В табл. 1 колонка τ показывает время, проведенное заявкой в соответствующем узле, также можно видеть, что $\alpha_i + \beta_i = 1$. Из графа вероятностей переходов получаем матрицу интенсивностей переходов процесса $X(t)$, которая, в случае наличия в системе N пользователей, имеет вид

$$\mathbf{A} = (a_{ij}) = \begin{cases} \frac{i_m \alpha_m}{\tau_m}, & \mathbf{j} = \mathbf{i} + \mathbf{e}_m, \\ \frac{i_m \beta_m}{\tau_m}, & \mathbf{j} = \mathbf{i} - \mathbf{e}_m, \\ -\sum_{m=1}^M \frac{i_m}{\tau_m}, & \mathbf{j} = \mathbf{i}, \\ 0 & \text{в остальных случаях,} \end{cases} \quad m = 1, \dots, M, \quad i, j \in X.$$

Таблица 1
Значения переходных вероятностей процесса $X(t)$ для модели движения пользователей

i	β	α	τ
1	0	1	16,709408
2	0,5044487	0,4955512	2,0139282
3	0,4931353	0,5068646	1,9840439
4	1	0	1,6421568

Введем процесс $F(t) = \int_0^t c(x)dx$, где $c(t)$ — скорость передачи данных в момент t . Обозначим через $T(x)$ момент первого достижения процессом $F(t)$ уровня x , $T(x) = \inf\{t | F(t) \geq x\}$, $\mathbf{Y}(x) = \mathbf{X}(T(x))$ и заметим, что неравенство $F(t) \leq x$ равносильно неравенству $T(x) \geq t$. Таким образом, процесс $(\mathbf{Y}(x), \mathbf{T}(x))$ является процессом с независимыми приращениями, заданным на цепи Маркова и является двойственным процессу $(\mathbf{X}(t), F(t))$ [2].

Обозначим через B матрицу интенсивностей переходов процесса $\mathbf{Y}(x)$. Его матрица переходных вероятностей задается формулой $\mathbf{Q}(x) = e^{\mathbf{R}^{-1}\mathbf{A}x}$ [2]. Процессы $\mathbf{X}(t)$ и $\mathbf{Y}(x)$ имеют одно и то же начальное распределение \mathbf{p} . Поэтому формула для первого центрального момента $w(x)$ достижения $T(x)$ уровня x $w(x) = \mathbf{p}\mathbf{Q}(x)\mathbf{R}^{-1}\mathbf{1}$ Здесь в качестве начального распределения \mathbf{p} мы берем стационарное распределение вероятностей, полученное решением СУР:

$$\begin{cases} \mathbf{A}\mathbf{p} = \mathbf{0}, \\ \mathbf{p}\mathbf{1} = 1. \end{cases}$$

Тогда для четырех крайних CQI и скоростях, взятых из табл. 2, можно получить графики среднего времени передачи данных для различных полос частот (5, 10, 15 МГц) и размера загрузочного блока, значение которого достигает до 50 МБ. Таким образом, соответствующие скорости рассчитываются как $r_i = c_i \cdot b$, где c_i — это скорость из табл. 2, a b соответствует рассматриваемой ширине полосы. Значения скоростей, приходящихся на 1 Гц, взяты из [3].

Таблица 2

Скорости передачи данных при различных уровнях качества

уровень качества	1	2	3	4
Скорость передачи данных (bit/s/Hz)	3,9023	4,5234	5,1152	5,5547

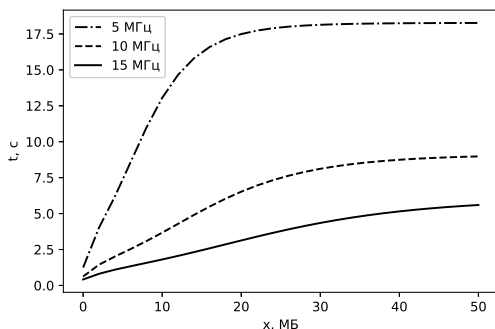


Рис. 2. Среднее время передачи данных

3. Заключение

Была построена модель для анализа среднего времени передачи данных в сети LTE с учетом перемещения абонентов внутри соты, проведен анализ среднего времени передачи данных в зависимости от объема передаваемых данных. В дальнейшем планируются рассмотреть большее число CQI, а также другие модели движения.

Благодарности

Публикация подготовлена при финансовой поддержке РФФИ в рамках научных проектов № 16-07-00766, 18-07-00567.

Литература

1. Кельберт М. Я., Сухов Ю. М. Вероятность и статистика в примерах и задачах. — МЦНМО, 2010. — Т. II: Марковские цепи как отправная точка теории случайных процессов и их приложения.
2. Наумов В. А., Мокров Е. В., Самуйлов К. Е. Анализ временных характеристик процесса передачи данных подвижным пользователям в сети LTE. — 2017. — № 4. — С. 79–84.
3. 3GPP TS 25.214 V8.9.0. 3rd Generation Partnership Project. Technical Specification Group Radio Access Network. Physical layer procedures (FDD). Release 8. — March 2010.

UDC 004.4

Mathematical Model of Average Data Transfer Time for Mobile Users in LTE Network

V. A. Razgonyaev*, E. V. Mokrov*, K. E. Samouylov*[†]

** Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

*[†] Institute of Informatics Problems, FRC CSC RAS
44-2 Vavilova Str., Moscow 119333, Russian Federation*

Email: vladislav.razgonyaev@gmail.com, mokrov_ev@rudn.university, samuylov_ke@rudn.university

in this study we consider a mathematical data transmission model for the mobile users moving within a cell. The cell is divided into zones that have different quality of service (channel quality indicator, CQI). The further the CQI from the center of the cell, the lower the data rate. The goal of the study is to analyze the average transmission time, depending on the volume of data transferred to mobile users.

Key words and phrases: multicasting, heterogeneous wireless network, service quality of service, markov control processes, motion models, average transmission time, dual process, queuing network.

УДК 621.391 + 621.37

Определение дальности гарантируемой радиосвязи в беспроводных телекоммуникационных сетях стандарта IEEE 802.11 с использованием программы ping

О. Н. Ромашкова*, В. Е. Самойлов†

* Кафедра прикладной информатики

ГАОУ ВО «Московский городской педагогический университет (МГПУ)»
2-ой Сельскохозяйственный проезд, д. 4, к. 1, Москва, Россия, 129226

† Кафедра «Автоматика телемеханика и связь на железнодорожном транспорте»
ФГБОУ ВПО «Российский университет транспорта (РУТ (МИИТ))»
ул. Образцова, д. 9, стр. 9, Москва, Россия, 127994

Email: ox-rom@yandex.ru, samoilov.1992@list.ru

В статье рассматриваются вопросы качества передачи данных в беспроводных телекоммуникационных сетях стандарта IEEE 802.11. В частности уделяется внимание определению дальности гарантируемой радиосвязи и вероятности потери пакетов. В качестве методики определения дальности гарантируемой радиосвязи и вероятности потери пакетов для маршрутизаторов стандарта IEEE 802.11 предлагается использовать программу ping. Кроме того, в статье приводятся результаты сравнения оценки качества передачи данных полученной по предлагаемой методике с данными технической документации на оборудование.

Ключевые слова: дальность гарантируемой радиосвязи, беспроводные телекоммуникационные сети, вероятность потери пакетов.

1. Введение

При построении и эксплуатации локальных беспроводных сетей стандарта IEEE 802.11 нередко приходится тестировать технику с целью определения качества услуг доступа к сети на том или ином расстоянии от точки доступа [1, 2, 5]. Тестирование беспроводного маршрутизатора может быть как предварительное (на лабораторном стенде), так и текущее (в процессе эксплуатации сети) [3, 6]. В обоих случаях наиболее простым и быстрым решением является использование встроенной в ОС Windows программы ping.

2. Определение дальности гарантируемой радиосвязи

Для определения дальности гарантируемой радиосвязи с использованием программы ping необходимо взять два ноутбука и роутер, поддерживающие беспроводной стандарт передачи данных IEEE 802.11 b/g/n. Ноутбуки и роутер должны быть подключены согласно схеме на рис. 1.



Рис. 1. Схема подключения оборудования для проведения тестирования

Настройка роутера и создание беспроводной локальной сети осуществляется через веб-интерфейс роутера. Для этого необходимо подключить роутер к ноутбуку

с помощью кабеля стандарта RJ 45 и, используя браузер, подключиться к стандартному IP-адресу роутера. В нашем эксперименте мы будем использовать wi-fi роутер Asus Rt-N12 стандартный IP-адрес которого 192.168.1.1.

Далее необходимо подключить второй ноутбук к созданной локальной сети. Для этого включается wi-fi адаптер ноутбука, выбирается созданная сеть, вводится соответствующий логин и пароль. После подключения проверяется обмен пакетами в сети.

Теперь на первом ноутбуке необходимо соответствующим образом настроить утилиту ping. Для этого сочетанием клавиш win+г открывается окно задач «Выполнить», туда вводится команда «cmd», после чего открывается командное окно.

В открытое командное окно вводится команда для запуска утилиты ping. В нашем случае эта команда выглядит следующим образом: «ping [ip адрес роутера] -l 1000 -t». С помощью этой командой мы формируем пакеты объёмом 1000 байт, которые будут передаваться от ПК1 к ПК2 и обратно до тех пор, пока мы не введём команду Ctrl+C.

Программа работает в течение длительного времени, достаточного для полноценного определения вероятности потери пакетов. После чего программа останавливается сочетанием клавиш Ctrl+C. Для нашего эксперимента достаточная длительность работы программы – 20 мин. Иначе можно запустить программу для передачи 1000 пакетов. По окончании работы программы мы увидим сведения о количестве отправленных, полученных и потерянных пакетов, кроме того, мы узнаем среднюю величину задержки при передаче данных по каналу.

После окончания работы программы и получения вероятности потери пакетов для данной точки, расстояние между роутером и приёмным ноутбуком увеличивается. Тестирование повторяется для нескольких точек с целью получения среднего значения вероятности потери пакетов. Предполагается начать с расстояния в 5 м между ПК1 и ПК2 и увеличивать его на 5 м для более точного получения результата.

3. Результаты тестирования беспроводной локальной сети стандарта IEEE 802.11

В ходе тестирования была получена вероятность потери (см. рис. 2) пакетов в 6 точках на расстоянии 5 м, 10 м, 15 м, 20 м, 25 м и 30 м от передатчика (см. табл. 1).

```

Администратор: C:\Windows\system32\cmd.exe
Ответ от 192.168.1.168: число байт=1000 время=321мс TTL=128
Ответ от 192.168.1.168: число байт=1000 время=311мс TTL=128
Ответ от 192.168.1.168: число байт=1000 время=296мс TTL=128
Ответ от 192.168.1.168: число байт=1000 время=311мс TTL=128
Ответ от 192.168.1.168: число байт=1000 время=312мс TTL=128
Ответ от 192.168.1.168: число байт=1000 время=307мс TTL=128
Ответ от 192.168.1.168: число байт=1000 время=307мс TTL=128
Ответ от 192.168.1.168: число байт=1000 время=307мс TTL=128
Ответ от 192.168.1.168: число байт=1000 время=311мс TTL=128
Ответ от 192.168.1.168: число байт=1000 время=311мс TTL=128

Статистика Ping для 192.168.1.168:
    Пакеты: отправлено = 10, получено = 10, потеряно = 0
    (0% потерь)
    Среднее время времени-прохода в мс:
    Минимальное = 296, Максимальное = 1226 мс, Среднее = 338 мс
    Сортировка
    C:\>
  
```

Рис. 2. Результаты тестирования в одной точке

Для уточнения результатов получим средние значения приведённых ранее параметров: среднее время задержки по 6 опытам – 332,33 мс; среднее кол-во отправленных пакетов – 1207 шт.; среднее кол-во потерянных пакетов – 11; процент потерь по 6 опытам – 0,91%.

Таблица 1

Результаты тестирования беспроводной локальной сети стандарта IEEE 802.11

Расстояние между ПК1 и ПК2, м	Кол-во отправленных пакетов	Кол-во потерянных пакетов	Среднее время задержки приёма-передачи, мс
5	1168	19	330
10	1219	12	323
15	1306	3	338
20	1167	23	330
25	1179	3	334
30	1203	7	339

4. Заключение

Согласно заявленным в технической документации на маршрутизатор данным, wi-fi роутер Asus Rt-N12 в условиях прямой радиовидимости обеспечивает передачу данных с вероятностью потери пакетов 1-3% на расстояние до 50 м. Сравнивая параметры, приведённые в технической документации с параметрами, полученными в реальных условиях, можно говорить об их правдивости. Кроме того, сравнивая значения вероятности потери пакетов со значениями, предлагаемыми для IP-телефонии [3, 4], можно утверждать, что качество передачи речи по этому каналу будет иметь оценку «отлично».

Литература

1. Горелов Г. В., Самойлов В. Е., Маркачёв С. А., И. Л. Попов. Измерение параметров радиолинии диапазона 2,4 ГГц при исключении влияния многолучевого распространения // Современные проблемы науки и образования. — 2015. — №1.
2. Ромашкова О. Н., Самойлов В. Е. К определению качества пакетной передачи речи в сетях подвижной связи // Научоёмкие технологии в космических исследованиях Земли. — 2017. — Т. 9, №3. — С. 39–44.
3. Ромашкова О. Н., Дедова Е. В. Живучесть беспроводных сетей связи в условиях чрезвычайной ситуации // T-Comm: Телекоммуникации и транспорт. — 2014. — Т. 8, № 6. — С. 40–43.
4. Ромашкова О. Н., Яковлев Р. И. Анализ моделей и методов для оценки живучести инфокоммуникационных сетей в условиях чрезвычайных ситуаций // T-Comm: Телекоммуникации и транспорт. — 2012. — Т. 6, № 7. — С. 165–170.
5. Ромашкова О. Н., Самойлов В. Е., Попов И. Л. Быстрая оценка качества передачи речи по беспроводным сетям связи в условиях ЧС // 13-я международная научно-практическая конференция «Достижения современной науки», секция «Технические науки». — М.: Научный центр «Олимп», 2016. — С. 624–626.
6. Горелов Г. В., Ромашкова О. Н., Житнов А. А. Искажения энергетического спектра речевого сообщения при использовании технологии Voice Over Wi-Fi // Телекоммуникации. — 2011. — № 1. — С. 10–12.

UDC 621.391 + 621.37

Determination of the Range of the Guaranteed Radio Communication in Wireless Telecommunication Networks of IEEE 802.11 Standard with the Use of Ping Program

O. N. Romashkova*, V. E. Samoylov†

* *Department of "Applied Informatics"
Moscow City Pedagogical University (MGPU)
2nd Selskokhozyajstvennyj passage, 4/1, Moscow, Russia, 129226*

† *Department of "Automatic Equipment, Telemechanics
and communication on railway transport"
Russian University of transport (RUT (MIT))
Obraztsova str., 9/9, Moscow, Russia, 127994*

Email: ox-rom@yandex.ru, samoilov.1992@list.ru

The article discusses the quality of data transmission in wireless telecommunication networks with IEEE 802.11. In particular, it is paid much attention to determining the range of guaranteed radio communication and the probability of packet loss. It is proposed to use the ping program as a method of determining the range of guaranteed radio communication and the probability of packet loss for IEEE 802.11 routers. In addition, the article presents the results of comparison of the quality of data transmission evaluation obtained by the proposed method with the data of the equipment technical documentation.

Key words and phrases: the range of guaranteed radio communication, wireless telecommunication networks, the probability of losing packets.

УДК 519.872.5

Модель мобильности в виде СЕМО с «пассивным» узлом**В. Н. Савич***, **В. А. Бесчастный***, **Д. Ю. Острикова***, **И. А. Гудкова***[†]* *Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*[†] *Институт проблем информатики
Федеральный исследовательский центр «Информатика и управление» РАН
ул. Вавилова, д. 44, к. 2, Москва, Россия, 119333*

Email: valeria.savich16@gmail.com, beschastny_va@rudn.university, ostrikova_dyu@rudn.university, gudkova_ia@rudn.university

В данной работе рассматривается модель мобильности с «пассивным» узлом для анализа движения пользователей в LTE-сетях.

Ключевые слова: ADR, CQI, LTE, мобильность пользователей, мультимедиа.

1. Введение

В данной работе рассматривается сеть массового обслуживания с M «активными» узлами и дополнительным «пассивным» узлом. В этом случае будем говорить о мультимедиа — форма передачи данных, при которой адресом назначения сетевого пакета является многоадресная группа. Рассматриваемая сеть является замкнутой — это сеть с множеством узлов без источника и стока, в которой циркулирует постоянное число N однотипных заявок [1]. Исходя из распределения вероятностей для данной модели можно улучшить производительность сети. Таким образом мы можем оптимально сформировать подгруппы для достижения максимальной ADR (Aggregate Data Rate) среди всех пользователей в соте. Распределение вероятностей модели СеМО с «пассивным» узлом основывается на модели с «активными» узлами.

2. Постановка задачи

В мобильных сетях пользователи могут быть как «активными», то есть пользоваться сотовой связью в текущий момент времени, так и «пассивными» — не получать сигнала базовой станции. Таким образом абоненты постоянно перемещаются по узлам с различными вероятностями переходов между узлами. Ранее в статье [2] была рассмотрена сеть массового обслуживания только с «активными» пользователями, то есть абоненты, в данном случае, пользуются мобильной сетью постоянно. Таким образом, сеть с одним «пассивным» узлом, которая представляет для нас интерес, является расширением для данной модели. Схема модели с «пассивным» узлом, изображена на рис. 1.

На рис. 1 M — количество активных узлов, нулевым обозначен «пассивный узел», а N — общее количество заявок системе, как в активных, так и пассивном узле. Через $\theta_{(mm)}$, $m = \overline{1, M}$ обозначены вероятности переходов между соседними узлами, где находятся активные пользователи, а через $\theta_{(0m)}$, где $m = \overline{1, M}$ обозначены вероятности переходов от «пассивного» узла к активным. Среднее время, пока пользователь не перемещается из области m обозначено через a_m . Таким образом, для того, чтобы доказать, что данная физическая система существует, она должна находиться в одном из доступных ей состояний, то есть удовлетворять условию нормировки:

$$\sum_{m=1}^M \theta_{(0m)} = 1.$$

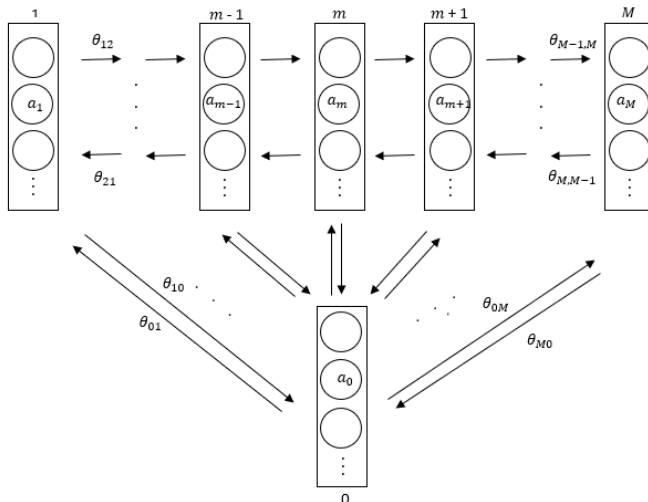


Рис. 1. Схема сети массового обслуживания с «пассивным» узлом

3. Заключение

В данной работе рассмотрена сеть массового обслуживания с «активными» узлами и одним «пассивным», выведено условие нормировки, удовлетворяющее условию существования данной системы. Модель может использоваться для анализа обслуживания мобильных пользователей в сети LTE.

Благодарности

Публикация подготовлена при поддержке программы РУДН «5-100» и при финансовой поддержке РФФИ в рамках научных проектов № 17-07-00142, 18-07-00567.

Литература

1. Башарин Г. П. Лекции по математической теории телеграфика. — М.: РУДН, 2009.
2. Ostrikova D., Rinaldi F., Beschastnyi V., Gudkova I., Militano L., Araniti G., Iera A., Samouylov K. Analytical Model for Multicast Subgrouping in 5G-Mobile eMBMS Environment // 9th International Congress “Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)”, 6–8 Nov. 2017. — IEEE, 2017. DOI: 10.1109/ICUMT.2017.8255188.

UDC 519.872.5

Model of Mobility as a Kind of Queuing Net with “Passive” Unit

V. N. Savich*, V. A. Beschastnyi*, D. Yu. Ostrikova*, I. A. Gudkova*[†]

** Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

*[†] Institute of Informatics Problems, FRC CSC RAS
44-2 Vavilova str., Moscow 119333, Russian Federation*

Email: valeria.savich16@gmail.com, beschastny_va@rudn.university, ostrikova_dyu@rudn.university, gudkova_ia@rudn.university

In this paper, we consider a mobility model with a “passive” node for analyzing user traffic in LTE networks.

Key words and phrases: ADR, CQI, LTE, mobility of users, multicasting.

**Прикладные информационные системы
и технологии**

Applied information systems and technologies

UDC 004.93'12

The System of Handwritten Letters Recognition Based on Neural Network

S. S. Volkov

*Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: volksergg@yandex.ru

This paper is about the method of pattern recognition using the example of handwritten letters of the English alphabet. A program was developed to solve this task. This program uses recognition algorithms based on neural networks. The paper studies several variants of the architecture of neural networks and their parameters. To train and test the effectiveness of the program, an open data set EMNIST (an extension of MNIST to handwritten letters) is used. The results of letters recognition on specific examples are demonstrated.

Key words and phrases: neural network, image recognition, handwritten letters.

1. Introduction

The task of recognizing handwritten symbols and texts in general is quite relevant nowadays. Translation of information from paper to electronic media will greatly simplify access to information resources. This work is devoted to the development of a program that should recognize the handwritten letters of the English alphabet. To implement this task, the python programming language was used, as well as libraries for this language, implementing deep neural networks (Keras, Theano). Training and analysis of neural network were performed on an open EMNIST data set.

2. Main section

The EMNIST data set contains a large number of samples of handwritten letters and numbers. Each symbol is contained in a separate image. We will consider a set containing only letters of the English alphabet. Image format:

- Size 28x28;
- Shades of gray (0-white color, 255 black);
- Images in binary form are recorded in four files (image file for learning network (124800 examples), file with marks for learning, image file for testing the quality of training (20,800 examples) and a file with labels for these images).

Examples of images are shown in picture (see Fig. 1).



Figure 1. Examples of EMNIST images

In the process of work, there were two programs developed. Each program implements its own neural network:

1. A fully connected neural network;
2. Convolutional neural network.

Let's consider both versions in more detail.

The first program implements a fully connected neural network for recognizing handwritten digits. The network has only two layers:

1. The input layer, 800 neurons. The value of each of the 784 image pixels arrives at the input of each neuron.
2. The output layer, 26 neurons. Each neuron corresponds to one of the classes of handwritten letters from A to Z.

The second program implements a convolutional neural network for recognizing handwritten letters. The training of this network will take longer, but the quality of recognition will be higher. We will use a network that consists of 6 layers:

1. Convolution layer, 75 feature maps, convolution kernel size: 5x5.
2. Pooling layer, size: 2x2.
3. Convolution layer, 100 feature maps, convolution kernel size: 5x5.
4. Pooling layer, size: 2x2.
5. Fully connected layer, 800 neurons.
6. Fully connected output layer, 26 neurons. Each neuron corresponds to one of the classes of handwritten letters from A to Z.

To teach both models, split the training data into two parts:

- Training sample, 99840 images;
- Validation sample, 24960 images.

The first model will be trained in 40 steps. The second model - in 10 steps. To test the quality of training models we use a test dataset that consists of 20,800 images. These images were not used in the training of neural networks.

The first model recognizes handwritten letters from the test data set with an accuracy of 84.81%. The second model trained 19 times longer but it recognizes handwritten letters from test data set with an accuracy of 94.61%

3. Conclusions

As a result of solving task of pattern recognition two programs were developed. Each program creates and trains a neural network that is able to recognize a handwritten letter with a high enough accuracy.

References

1. Castro-Bleda M. J., Espana-Boquera S., Zamora-Martinez F. Handwritten Text Recognition. — Zaragoza, 2012.
2. Govindaraju V., Srihari R., Srihari S. Handwritten text recognition. Document Analysis Systems, 1995.
3. Kim G., Govindaraju V., Srihari S. N. An architecture for handwritten text recognition systems // International Journal on Document Analysis and Recognition. — 1999. — T. 2. — N. 1.
4. Kulik S. D. Neural network model of artificial intelligence for handwriting recognition // Journal of Theoretical & Applied Information Technology. — 2015. — T. 73. — N. 2.
5. Marti U. V., Bunke H. On the influence of vocabulary size and language models in unconstrained handwritten text recognition // Document Analysis and Recognition, 2001. Proceedings. Sixth International Conference on. — IEEE, 2001.
6. Zimmermann M., Bunke H. Optimizing the integration of a statistical language model in HMM based offline handwritten text recognition // Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on. — IEEE, 2004. — T. 2.

УДК 004.93'12

Система распознавания рукописных букв на основе нейронной сети

С. С. Волков

*Кафедра информационных технологий,
Российский университет дружбы народов,
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: volkserrg@yandex.ru

В статье предложен метод распознавания образов на примере рукописных букв английского алфавита. Для решения задачи используется программа, реализующая алгоритмы распознавания на основе нейронных сетей. В статье рассмотрены несколько вариантов архитектуры нейронных сетей и их параметры. Для обучения и проверки эффективности работы программы используется открытый набор данных EMNIST (an extension of MNIST to handwritten letters). Продемонстрированы результаты распознавания символов на конкретных примерах.

Ключевые слова: нейронная сеть, распознавание образов, рукописные буквы.

UDC 614.2:519.237.8

Prospects of Cluster Analysis Application for Reengineering of Clinical Business Processes

T. V. Grigoryeva, V. V. Zhukov

*Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: tgrigoryeva97@gmail.com, zhukov_vv@rudn.university

The purpose of this study is the consideration of the possibility of implementation of cluster analysis for control over the flow of medical processes automation and simplification. For this, firstly, the task of such analysis was formulated and result to be achieved by the solution was determined. Secondly, the need for preliminary factor analysis was put under question, and the features of this method application to business processes, in particular, to the medical ones, were highlighted. In addition, the description of the approximate structure of the medical organization was given with the process versus hierarchical approach to management. The importance of the preliminary division of processes by the direction of activity is emphasized and cases in which the application of such analysis is considered appropriate are underlined. Moreover, the readers' attention is focused on the need for a number of cross-cutting characteristics of business processes. In conclusion, it is noted that this type of analysis can become a part of actions that provides up-to-date information about the system of clinical processes flow and enables to carry out their timely reengineering.

Key words and phrases: cluster analysis, process approach to the management, clinical business intelligence, reengineering of business processes.

1. Introduction

In spite of the achievements of modern medicine and healthcare, in recent years there has been a significant increase in attention to health issues at the national level [1]. In particular, people are concerned about the access to primary health care in polyclinics and on an outpatient facilities. In many ways, this problem is associated with outdated methods of liaising with patients and absence of patient flows distribution systems that leads to creation of non-medical barriers to accessing health care services.

In some regions of the Russian Federation "Lean polyclinic" is currently being tested, designed to save time for patients and staff and to improve the efficiency of medical facilities. Through clinical business analysis, the program optimizes the functioning of polyclinics and outpatient departments, in particular, registries (receptions). The term "clinical business intelligence" refers to the use of data analysis to improve care delivery. This concept includes a wide range of information technologies that gather data created during the delivery of medical care and analyze that clinical data to help managers and clinicians make better decisions [2].

The operation of such systems generates a large amount of data that have to be analyzed correctly and quickly. This article raises the issue of using cluster analysis methods to automate the medical business processes monitoring.

2. Background and results

Cluster analysis is a method of multidimensional classification based on the results of separate observations representation in the form of points of a suitable geometric space, followed by the segregation of groups of these points as "clots" (clusters, taxons). This method of research gained popularity alongside the possibility of automated processing of large databases and data warehouses [3].

With the help of cluster analysis of business processes, the search for their natural division into groups will be conducted with the most similar characteristics.

Let the set of processes $P = \{P_1, P_2, \dots, P_n\}$ consists of n processes. Each process P_j has a range of characteristics (for example, process execution time, patient waiting time, number of executors), the quantitative value each i th of which is designated as x_{ij} . A vector $X_j = [x_{ij}]$ contains the data of all dimensions of each process. The result is a set of vectors $X = \{X_1, X_2, \dots, X_n\}$, which characterizes the entire set of processes.

As a result of calculating the distance between points of set X , we obtain proximity matrices or matrices of distances between objects (processes). Under the task of cluster analysis of business processes, we will mean the need to divide the set of processes P into a number of m subsets, in other words, clusters, which is much smaller than the original number of processes n . An important condition is that each process belongs to only one cluster while and objects in different clusters differ considerably. Inside the sets, the processes are, on the contrary, as similar as possible in accordance with the proximity matrices [3].

As a solution to business processes cluster analysis task, we will achieve some division, illustrating particular consistency among the processes, for example, similar problems or, in contrast, the successful delivery of wide range of processes.

It is important to note that the number of clusters m should be significantly less than the initial number of processes, otherwise the method will not have any practical significance.

Moreover, it will be reasonable to use factor analysis for data reduction before cluster analysis, because there is a possibility that the analyst will receive either similar data from similar processes or different experimental measurements of the same process as the source data.

The implementation of a process approach to the management of a medical organization implies the transformation from a classical hierarchical management system to a health facility as a set of interconnected and interdependent business processes with a broad delegation of responsibilities to the executor [4]. However, there is no need to perform cluster analysis of all processes of one organization as a whole.

Three process groups can be distinguished in the construction of the medical institution process model: management support (operational management, strategic planning, quality control), medical activities support and auxiliary activity support (accounting, legal assistance, HR, etc.) [5].

Consequently, it seems reasonable to perform cluster analysis within the same group of processes or within different measurements of the same process (for example, to identify external factors interfering the correct process delivery).

In addition, the business process can be monitored and controlled only if there is system of their measurement indicators. Based on these data, health facility managers will be able to take preventive or corrective actions and carry out analytical actions [6]. However, the usage of cluster analysis requires not only existence of indicating characteristics, but their consistency as well. This method does not include the analysis of completely dissimilar features, so cluster clinical business analysis requires cross-cutting characteristics with a defined measurement system in addition to the indicators typical for any process. The simplest of such features can be the patient's satisfaction with medical care, however, it is not the most indicative criterion, unlike, for example, the time and laboriousness.

3. Conclusions

The current health care system in Russia requires revision, taking into account the growing discontent with the medical services quality. The State Corporation "Rosatom" along with the Ministry of Health of the Russian Federation introduces a process approach to the management of medical institutions in a number of regions within the framework of a pilot project. The first positive results have already been achieved. However, for full-fledged conclusions about the effectiveness of the work and analysis, it is necessary to implement a set of actions that provide up-to-date information on system operation and enable to carry out timely reengineering of business processes. One of

such measures may be the cluster analysis, through which one can improve and simplify control over the work of medical institution by dividing processes into groups (clusters) based on salient problems.

References

1. WCIOM, Пресс-выпуск №3277. Проблемный фон страны: итоги года , (Russian), [Press release No. 3277. Problematic background of the country: results of the year], 2016.
2. What Is Clinical Business Intelligence? (Module 1), HIMSS, 2013.
3. D. Kuznetsov, T. Troshina, Кластерный анализ и его применение , (Russian), [Cluster analysis and its applications], Ярославский педагогический вестник, no. 4, 2006, 103-107.
4. S. Ermakova, E. Evdokimova, К вопросу о процессно-ориентированной модернизации системы управления в медицинских организациях , (Russian), [On the issue of process-oriented modernization of the management system in medical organizations], Вектор науки Толятинского государственного университета, no. 4(23), 2013, 96-99.
5. D. Belyshev, Y. Guliev, V. Malykh, Моделирование бизнес-процессов медицинской организации (лечебно-профилактического учреждения) , (Russian), [Modeling of business processes of medical organization (treatment-and-prophylactic institution)], Врач и информационные технологии вестник , no. 5, 2014, 78-90.
6. P. Gonçalves, M. Hagenbeek, J. Vissers, Hospital process orientation from an operations management perspective: development of a measurement tool and practical testing in three ophthalmic practices, BMC Health Services Research, 2013 13:475. doi:10.1186/1472-6963-13-475.

УДК 614.2:519.237.8

Перспективы применения кластерного анализа для реинжиниринга клинических бизнес-процессов

Т. В. Григорьева, В. В. Жуков

*Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: tgrigoryeva97@gmail.com, zhukov_vv@rudn.university

Целью данной статьи было рассмотрение возможности использования методов кластерного анализа статистических данных для автоматизации и упрощения контроля за функционированием медицинских процессов. Для этого, во-первых, была сформулирована задача подобного анализа и определен достигаемый при решении результат. Во-вторых, поднят вопрос предварительного факторного анализа и выделены особенности применения данного метода к бизнес-процессам, в частности медицинским. Кроме того, дано описание примерной структуры медицинской организации с процессным подходом к управлению в отличие от иерархического. Подчеркнута важность предварительного разделения процессов по направлению деятельности и выделены случаи, в которых применение подобного анализа рассматривается целесообразным. Дополнительно сделан акцент на необходимость ряда сквозных характеристик бизнес-процессов. В заключение отмечено, что данный вид анализа может стать частью комплекса мер, используемых для получения актуальной информацию о функционировании системы клинических процессов и позволяющий проводить их своевременный реинжиниринг.

Ключевые слова: кластерный анализ, процессный подход к управлению, клинический бизнес-анализ, реинжиниринг бизнес-процессов.

UDC 519.248.6:004.855.5:57.084.1:616-006.66

Application of Support Vector Machine to Predict the Onset of Breast Cancer in Mouse Model

V. V. Zhukov*, D. A. Aronov†, S. G. Semushina†, E. V. Moiseeva†

* *Department of Information Technologies*

*Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

† *Laboratory of Biotechnology*

*Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry
16/10 Miklukho-Maklaya str., Moscow, 117997, Russian Federation*

Email: zhukov_vv@rudn.university, evmoise@gmail.com

Early detection of breast cancer (BC) is still one of the most urgent issues of both clinical and experimental oncology. Currently, mammography can be considered as the only existing mass screening technique. Nevertheless, a number of significant drawbacks of this method impose restrictions on its use and raises the issue of developing fundamentally new alternative technologies for early diagnosis and prediction of the occurrence of BC. Earlier, we demonstrated that in a population of mice with the high incidence rate of naturally occurring mammary cancer (MC) immunological parameters of blood changed long before the tumor onset. In the current study, we proposed a mathematical model to predict the MC occurrence in mice, based on the previously measured immunological parameters of the blood. We compared the efficiency of different classification algorithms; the best results were obtained for Support Vector Machine classifier with the accuracy of 72.16%, which could be considered as an acceptable preliminary result for a limited sample size. Thus, several immunological parameters of the blood apparently could be used as the biomarkers for early diagnosis of breast cancer.

Key words and phrases: breast cancer prediction, statistical classification, support vector machine, mouse model.

1. Introduction

Breast cancer (BC) is the most common oncologic disease among Russian women (21%); annual incidence and mortality levels of BC remain high [1]. Early detection of BC is essential for selecting the most efficient individualized therapeutic strategy, which allows to significantly increase the life expectancy of a patient [2]. Currently, mammography is the most common procedure for mass population screening of the female population for the detection of breast cancer.

Women older than 40 years are recommended to undergo mammography annually [3]. This procedure often leads to overdiagnosis regardless of patient's age and may result in excessive treatment with many side effects [4]. In addition, it was clearly reported that frequent mammograms could provoke early tumor manifestation themselves, especially in women with a predisposition to BC [5]. Mathematical models for predicting the BC onset could serve as an alternative to mammography. In 1989, M. Gale proposed a model that calculates the individual risk of BC using the data from anamnesis of a woman [6]. However, it is not possible to predict the time of tumor appearance after passing the test by the patient.

Aged mouse females with the high incidence of naturally occurring mammary cancer (MC) that reproduce pathomorphology of some human BC types may serve as a good model for human familial BC [7]. Earlier, we demonstrated predictive value of some routine immunological blood parameters that were changed long before the tumor onset in a mouse model of BC [8]. The aim of the current study was to develop a mathematical model of individual prediction of the MC outcome, using the immunological blood parameters measured in aging mouse females long time before MC onset.

2. Main section

We used a population of intact aging non-SPF females ($n = 81$) of two genetically related mouse strains BLRB-Rb(8.17)11em and BYRB-Rb(8.17)11em characterized with the high incidence of naturally occurring MC. We collected 200 μ l of blood from retroorbital venous sinus and followed all mice until death. The following immunological parameters of the blood were measured with flow cytometry: lymphocytes, monocytes, neutrophils, eosinophils, T-lymphocytes, T-helpers, cytotoxic T-lymphocytes, B-lymphocytes, NK-cells, NKT-cells, activated T-helpers. Health status, newly appeared MC, and survival of mice were monitored after blood samples collection until death. After the end of experiment, mice were stratified into two classes according to the outcome: class 0 ($n = 55 / 81$; 67.9 %) – females did not have symptoms of MC; class 1 ($n = 26 / 81$; 32.1 %) – females with detected MC.

We used STATISTICA 12 and Python 2.7.14 with the Scikit-learn library [9] for data analysis and developing the classification models. At the first step all the data were split into training ($n = 57 / 81$; 70.4 %) and test ($n = 24 / 81$; 29.6 %) set. The division of the dataset was performed in a stratified fashion to maintain the ratio between the initial classes in the training and test samples. Exploratory analysis of the data was performed to determine the most significant features and reduce the dimensionality of data. Fisher Score [10] was calculated for each parameter, then they were ranked in order of importance. The results of the exploratory analysis confirmed the hypothesis of the immune system aging (immunosenescence) and multidirectional changes in immunological characteristics of mice as the age. The final group of informative parameters was compiled with the following variables: age at the blood collection point, levels of neutrophils, eosinophils, B-cells, NK-cells, NKT-cells.

Stratified 9-fold cross-validation over the training set was used to determine the optimal model and its hyperparameters, compare algorithms and their settings. To compare the efficiency of algorithms, we used the F_1 score (the harmonic average of the precision and recall). Among several classification methods (such as k-nearest neighbors, support vector machine, multilayer perceptron, decision tree and naive Bayes classifier) and various variants of the values of their hyperparameters, SVM classifier with RBF-kernel showed the greatest efficiency.

We used repeated stratified holdout validation with the same size of the test sample (24 cases) to determine the average values of the algorithm's efficiency metrics. The average accuracy of RBF-kernel SVM classifier on test set was 72.16% ($\pm 0.05\%$); the average sensitivity was 59.75% ($\pm 0.12\%$); the average specificity was 78.4% ($\pm 0.07\%$). This result could be considered as an acceptable preliminary result for a limited sample size, however, further analysis of the errors in classification is necessary to improve the work of this algorithm.

3. Conclusions

We used murine female population with the predisposition to MC to develop Support Vector Machine based on the immunological blood parameters. The individual onset of murine mammary cancer was predicted with an accuracy of 72.16%. Therefore, even with a limited sample size ($n = 81$), it is possible to achieve an acceptable result for determining the cancer onset. In conclusion, the immunological parameters of the blood apparently could be used as the surrogate biomarkers for predicting the breast cancer occurrence, at least for women with a familial predisposition along with other diagnostic procedures.

References

1. P. E. Goss, K. Strasser-Weippl, B. L. Lee-Bychkovsky, et al., Challenges to effective cancer control in China, India, and Russia, *The Lancet Oncology* 15 (5) (2014) 489–538.

2. S. Walters, C. Maringe, J. Butler, et al., Breast cancer survival and stage at diagnosis in Australia, Canada, Denmark, Norway, Sweden and the UK, 2000-2007: a population-based study, *British Journal of Cancer* 108 (5) (2013) 1195–1208.
3. P. C. Gøtzsche, M. Nielsen, Screening for breast cancer with mammography, *The Cochrane database of systematic reviews* 1 (2011) CD001877.
4. M. Løberg, M. L. Lousdal, M. Bretthauer, M. Kalager, Benefits and harms of mammography screening, *Breast Cancer Research* 17 (1) (2015) 63.
5. M. C. Jansen-van der Weide, M. Greuter, L. Jansen, et al., Exposure to low-dose radiation and the risk of breast cancer among women with a familial or genetic predisposition: a meta-analysis, *European Radiology* 20 (11) (2010) 2547–2556.
6. M. H. Gail, L. A. Brinton, D. P. Byar, et al., Projecting individualized probabilities of developing breast cancer for white females who are being examined annually, *Journal of National Cancer Institute* 81 (24) (1989) 1879–1886.
7. E. Moiseeva, Anti-breast cancer drug testing. Original Approaches. Novel Set of Mouse Models, Lambert Academic Publishing, Saarbrücken, Germany, 2009.
8. D. A. Aronov, S. G. Semushina, E. V. Moiseeva, Prognostic value of immunological blood parameters in the mouse model of familial breast cancer, *Modern problems of science and education* 5 (2017).
9. F. Pedregosa, G. Varoquaux, A. Gramfort, et al., Scikit-learn: Machine learning in Python, *Journal of Machine Learning Research* 12 (2011) 2825–2830.
10. R. O. Duda, P. E. Hart, D. G. Stork, *Pattern Classification*, 2nd Edition, John Wiley & Sons, 2012.

УДК 519.248.6:004.855.5:57.084.1:616-006.66

Применение машины опорных векторов для предсказания возникновения рака молочной железы в мышиной модели

В. В. Жуков*, Д. А. Аронов†, С. Г. Семушина†, Е. В. Моисеева†

* Кафедра информационных технологий,
Российский университет дружбы народов,
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198

† Лаборатория биотехнологии,
Институт биоорганической химии им. Шемякина-Овчинникова,
ул. Миклуто-Маклая, д. 16/10, Москва, Россия, 117997

Email: zhukov_vv@rudn.university, evmoise@gmail.com

Ранняя диагностика рака молочной железы (РМЖ) по-прежнему остается одной из наиболее актуальных задач клинической и экспериментальной онкологии. На сегодняшний день маммография является единственным методом массового скрининга населения. Однако набор существенных недостатков, присущих этому методу, накладывает ограничения на его использование и поднимает проблему разработки фундаментально новых технологий для ранней диагностики и предсказания возникновения РМЖ. Ранее мы показали, что в популяции мышей с высокой частотой естественно возникающего РМЖ иммунологические параметры крови изменялись задолго до возникновения опухоли. В данной работе мы сформировали математическую модель для предсказания возникновения РМЖ в мышиной модели на основе предварительно измеренных иммунологических показателей крови. Мы сравнили эффективность различных алгоритмов классификации; лучшие результаты были показаны классификатором на основе машины опорных векторов с точностью 72,16%. Такой результат можно считать приемлемым для выборки ограниченного размера. Таким образом, можно говорить о том, что некоторые иммунологические параметры могут быть использованы как биомаркеры для ранней диагностики рака молочной железы.

Ключевые слова: предсказание рака молочной железы, статистическая классификация, машина поддерживающих векторов, мышиная модель.

УДК 004.4

Разработка метода поддержки принятия решений о выборе тактики хирургического лечения грыжи диска

М. О. Зубрихина*[†], А. И. Молодченков*[‡]

** Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

*† ООО Технологии системного анализа
проспект 60-летия октября, д. 9, Москва, Россия, 117312*

*‡ Лаборатория «Интеллектуальные технологии и системы»
Федеральный исследовательский центр «Информатика и управление» РАН
проспект 60-летия октября, д. 9, Москва, Россия, 117312*

Email: m.zubrikhina2014@yandex.ru, aim@tesyan.ru

В медицине на данный момент накоплено большое количество информации о результатах лечения заболеваний позвоночника различными консервативными и оперативными методами. Сложность и многочисленность имеющихся в арсенале у нейрохирурга возможностей декомпрессии и фиксации позвоночника, разнообразие и индивидуальность вариантов течения заболеваний, многочисленность анализируемых факторов и симптомов, приближенность оценивания результатов требует привлечения математических методов и компьютерных средств, которые позволяют анализировать имеющийся формализованный материал и результативно использовать его для поддержки принятия решений при диагностике и лечении конкретного пациента. В статье описан метод поддержки принятия решений о выборе тактики хирургического лечения грыжи диска.

Ключевые слова: компьютерные науки, информационные технологии, искусственный интеллект, машинное обучение.

1. Введение

При лечении дегенеративных заболеваний позвоночника, в частности грыжи диска, перед врачом всегда стоит задача выбора метода лечения. Сам выбор метода зависит от результатов предварительного обследования больного перед хирургическим вмешательством и оказывает влияние на исход лечения. Ошибка при выборе метода вмешательства может привести к инвалидности.

Эта работа посвящена разработке метода поддержки принятия решений о выборе тактики хирургического лечения грыжи диска. Для разработки метода поддержки принятия решений о выборе тактики хирургического лечения грыжи диска были предоставлены данные о 303 пациентах центром нейрохирургии имени Н. Н. Бурденко. Среди всех пациентов 133 пациента были прооперированы с помощью микродиссектомии, 34 — стабилизации и 136 — с помощью радиочастотной денервации. Информация о каждом пациенте описывалась 28 параметрами, которые включают в себя рост (в сантиметрах), вес (в килограммах), пол, длительность заболевания (количество лет), длительность настоящего обострения (в неделях), наличие операций на позвоночнике в анамнезе и др. Значения клинических и рентгенологических показателей на каждого пациента записывались по трех-балльной шкале, в зависимости от степени их выраженности. Например, параметр «Боль в поясничной области» принимает следующие значения: «Нет», «Иногда», «Постоянная», а параметр «Боль в ногах» — «Нет», «В одной ноге», «В обеих ногах». Кроме того, для каждого пациента анализировались антропометрические и анамнестические показатели — рост, вес, пол, длительность заболевания, длительность настоящего обострения, наличие операций на позвоночнике в анамнезе.

2. Описание метода

Для разработки метода поддержки принятия решений о выборе тактики хирургического лечения пациентов с грыжей диска необходимо построить три предсказательные модели для каждой из хирургических тактик, а именно микродискектомии, стабилизации и радиочастотной денервации позвоночных суставов.

Для построения моделей первым шагом необходимо было произвести очистку, конструирование и стандартизацию признаков. Проблемой многих медицинских данных является несбалансированность выборки, то есть количество положительных примеров намного больше количества отрицательных примеров. Для работы с такими данными существуют методы, которые позволяют искусственным образом добавить или удалить примеры. Так как в данной задаче обучающая выборка является очень небольшой, что в целом часто приводит к переобучению моделей, то для приведение ее к сбалансированному виду необходимо использовать методы, которые могут искусственно добавить примеры. Для работы с несбалансированными данными в этой работе был применен алгоритм SMOTE [1]. Далее были выбраны наиболее лучшие модели предсказывающие результат применения каждой из хирургических тактик и вычислена точность их работы. Данными моделями является логистическая регрессия [2], случайный лес [3] и адаптивный бустинг [4]. В качестве метрики качества для оценки моделей предсказывающих исход каждой из хирургических тактик была использована F1-мера. Как видно из табл. 1, лучшее качество дает алгоритм случайный лес, а худшее логистическая регрессия.

Таблица 1
F1-мера моделей предсказывающих результат применения хирургических тактик

Название алгоритма	РЧД	Стабилизация с декомпрессией	Микродискектомия
Случайный лес	0.94	0.82	0.98
Адаптивный бустинг	0.9	0.735	0.93
Логистическая регрессия	0.85	0.70	0.63

Метод поддержки принятия решений о выборе тактики хирургического лечения грыжи диска состоит из нескольких шагов. Сначала вводится информация о состоянии пациента. Далее система на основе вышеописанных прогностических моделей анализирует полученные данные, после чего врач получает информацию о вероятностях исходов при различных хирургических вмешательствах. На основе данной информации врач принимает решение о выборе лечения.

3. Заключение

В работе описан метод поддержки принятия решений о выборе тактики хирургического лечения грыжи диска. Представлено описание алгоритмов, которые позволили построить прогностические модели, показавшие лучшее качество. К данным алгоритмам относятся адаптивный бустинг и случайный лес.

Результаты экспериментальных исследований показали, что построенный метод поддержки принятия решений о выборе тактики хирургического лечения хорошего качества и может быть применен на практике. На его основе было разработано веб-приложение, которое позволяет в реальном времени ввести данные для анализа пациента и получить рекомендацию об успешности того или иного хирургического вмешательства, которое является отдельным модулем Российского вертебрологического регистра.

Литература

1. Chawla N. V. et al. SMOTE: synthetic minority over-sampling technique // Journal of artificial intelligence research. — 2002. — Vol. 16. — Pp. 321–357.
2. Harrell F. E. Ordinal logistic regression // Regression modeling strategies. — New York: Springer, 2001. — Pp. 331–343.
3. Liaw A. et al. Classification and regression by randomForest // R news. — 2002. — Vol. 2, no. 3. — Pp. 18–22.
4. Freund Y., Schapire R., Abe N. A short introduction to boosting // Journal-Japanese Society For Artificial Intelligence. — 1999. — Vol. 14. — Pp. 771–780.

UDC 004.4

Development of the Method of Supporting Decision-Making about the Choice of Tactics of Surgical Treatment of a Herniated Disc

M. O. Zubrikhina*[†], A. I. Molodchenkov*[‡]

** Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

*† Systems Analysis Technology
9 Prospekt 60th Anniversary of October, Moscow, 117312, Russian Federation*

*‡ Intelligent Technologies and Systems Laboratory
Federal Research Center "Computer Science and Control" of PAS
9 Prospekt 60th Anniversary of October, Moscow, 117312, Russian Federation*

Email: m.zubrikhina2014@yandex.ru, aim@tesyan.ru

There is an intensive accumulation of information about the results of spine treatment in medicine. Complexity and multiplicity of the neurosurgeon's ability to decompress and fix the spine; diversity of the variants for treatment of diseases; multiplicity factors and symptoms for analyze requires the attraction of tools for analyze of existing formalized material and use it effectively to support Decision making in diagnosis and treatment. The article describes a method of supporting the decision-making on the choice of tactics of surgical treatment of a herniated disc.

Key words and phrases: computer science, information technologies, artificial intelligence, machine learning.

УДК 519.689

Пример применения OLAP

А. В. Катрич

*Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: arkrstm@gmail.com

В данной работе будет рассмотрено использование OLAP системы на примере риэлторской компании. Будет представлено упрощенное описание пошаговой настройки, которое в дальнейшем может быть использовано как руководство по самостоятельному развертыванию OLAP.

Ключевые слова: многомерные базы данных, аналитика, OLAP, многомерный куб.

1. Введение

На рынке информационных технологий уже долго время существует технология OLAP. Это набор инструментов и методов для оперативной аналитической обработки данных, включающий в себе организацию их хранения, сбора и анализа и использующий концепцию хранения данных в многомерных структурах. OLAP внедряется с целью увеличения скорости обработки больших объемов данных, сокращения временных и трудовых затрат на подготовку аналитической отчетности сотрудниками. Технология призвана заменить длительный процесс подготовки запросов в базы данных на единую систему, из которой каждый пользователь может получить необходимые данные, обладая при этом минимальными навыками программирования и работы с БД. Для увеличения скорости работы с данным используется предварительная агрегация, позволяющая вычислить большинство значений при построении куба и в дальнейшем вытаскивать их из хранилища.

Основными понятиями OLAP куба являются:

- 1) Измерение — свойство рассматриваемого объекта
- 2) Мера — значение, расположенное на пересечении измерений, количественно характеризующее объект
- 3) Иерархия — способ многоуровневой организации элементов измерений

В зависимости от числа измерений, выбранных значений, способа агрегации, уровня детализации возможно извлекать из куба разнообразные совокупности данных. Фиксация части измерений называется срезом куба и является одной из основных операций в OLAP [1].

2. Реализация OLAP

Рассмотреть упрощенную версию внедрения и использования OLAP можно на примере работы одной риэлторской компании. Предположим, что она достаточно крупная, имеет множество клиентов, как продавцов недвижимости, так и покупателей. Чтобы извлечь максимальную прибыль, компания постоянно анализирует результаты взаимодействий с тем или иным клиентом, с целью определения необходимости прекращения договора с ним, продления или даже поиска новой клиентской базы. Она так же динамически меняет цены на недвижимость, опираясь на статистику продаж. Руководство решает внедрить OLAP, с целью ускорить процесс сбора отчетов о деятельности компании.

Первым делом подготавливается новая структура БД в форме звезды или снежинки (рис. 1).

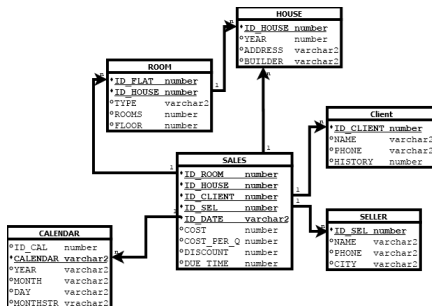


Рис. 1. Структура БД в виде звезды

В центре стоит таблица фактов, содержащая информацию о каждой сделке в компании. К ней по внешним ключам присоединены таблицы измерений, описывающие атрибуты сделки [2]. Затем с помощью программы Iccube, представляющей из себя OLAP сервер, выберем созданную базу как источник данных и зададим измерения, иерархии, меры, добавим вычисляемые меры, зададим агрегацию. Всего было создано 5 иерархий:

- 1) Год → Месяц → День
- 2) Город компании-продавца → Название компании-продавца
- 3) Название компании-клиента
- 4) Застройщик дома → Год постройки → Номер дома
- 5) Тип недвижимости(квартира, офис и пр.) → Номер квартиры

Меры и способы их агрегации:

- 1) Сумма сделки – количество
- 2) Цена за квадратный метр – среднее значение
- 3) Длительность сделки – максимальное значение
- 4) Доход компании риэлтора (вычисляемое значение) = Сумма сделки*0.15

Далее воспользуемся Microsoft Excel как средством доступа к данным в кубе, благодаря его возможности подключаться к внешним источникам и представлять извлеченные знания в виде сводной таблицы. Подготовим отчет, показывающий доход от сделок за 2016 и 2017 года в разрезе покупатель/продавец из Вологды (табл. 1).

Можно также построить диаграмму, показывающую динамику сделок с участием компании Ворон в 2018 году (рис. 2).

3. Заключение

В статье был разобран пример применения OLAP технологии на примере работы риэлторской компании. В результате настройки системы были подготовлены два варианта представления отчетов: в виде сводной таблицы и диаграммы. Полученные данные позволят руководству компании оценить результаты работы с определенными клиентами, а так же получить статистику совершенных сделок. Внедрение OLAP является ожидаемым решением там, где аналитикам необходим быстрый и эффективный инструмент построения отчетности. Полученные результаты и описанный процесс настройки системы могут быть использованы как руководство для тех, кто планирует разобраться в OLAP инструменте.

Таблица 1
Часть отчета за 2016 и 2017 года в разрезе покупатель/продавец

Profit	Вологда			
	Берега		Кобольд	
	2016	2017	2016	2017
Азбука	20528534	12195520	10249322	4881270
Альпы	8168767	11310546	10069692	16639557
Альфа	14296401	12607998	12032052	17915806
Байкал	13199466	12076648	18255141	6328680
Бетта	10354699	15589731	11400333	15283548

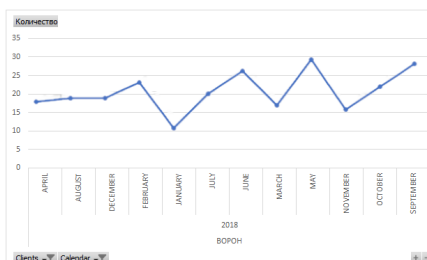


Рис. 2. Диаграмма динамики сделок

Литература

1. Барсегян А. А., Куприянов М. С., Степаненко В. В., Холод И. И. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP. — 2-е из., перераб. и доп. — СПб.: БХВ-Петербург, 2007. — 384 с.
2. Olap.ru — <http://olap.ru>.

UDC 519.689

The Example of Using OLAP

A. V. Katrich

*Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: arkrstm@gmail.com

In this paper we will review the use of the OLAP system on the example of real company. A simplified step-by-step configuration description will be provided, which can be used as an OLAP deployment guide.

Key words and phrases: multidimensional databases, analysis, OLAP, multidimensional cube.

UDC 519.248:616.895.8

Meta-Analysis of Medical Publications on Psychopathology: a Posterior Probability Calculation

A. N. Kozlovskii, V. V. Zhukov

*Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: arcozlovskij@yandex.ru, zhukov_vv@rudn.university

Mental health of a patient, being a rather complex system, requires qualified treatment by psychiatrists. However, in the conditions of limited time for admission and a large number of incoming patients, it may be difficult to render equally qualified assistance to everyone.

In the current study we underlined the application of Bayes theorem in diagnostics of psychopathologies, calculation of a posterior probability of diagnosis of schizophrenia, based on the presence of one or more types of hallucinations in a patient, and an assessment of the impact rates of hallucinations types. Source data were derived based on a meta-analysis of medical publications. The highest posterior probability of schizophrenia was observed in case of combined auditory and visual hallucinations (62.78%); the smallest posterior probability was observed in case of observing tactile hallucinations (3.48%).

The results obtained in this study can be used in diagnostics or patient prioritization, which could significantly reduce the burden on psychiatrists.

Key words and phrases: hallucinations, schizophrenia, psychopathology, Bayes theorem, meta-analysis.

1. Introduction

Mental health is a critical component of human health [1]. Early detection, prevention and treatment of mental disorders play an important role in maintaining a favorable picture of mental health of the population. One of the important tasks facing psychopathologists is the formulation of a correct diagnosis. Currently this problem is being solved by usage of criteria, described in DSM-5 and ICD-10 [2], with subsequent monitoring of mental health in clinics. However, the question of the influence of various symptoms on the probability of diagnosing a mental disorder remains open. Bayes theorem is a reasonable instrument for solving this problem; it becomes possible to illustrate the changes in probabilities of having a disease depending on the signs observed using statistical data.

Schizophrenia is considered in this study as a research object. Several types of hallucinations, one of the key diagnostic indicators of schizophrenia, were selected as registered signs. Thus, calculating the posterior probability of schizophrenia under the condition of observing a certain type of hallucinations, it is possible to estimate the risk of schizophrenia in a patient.

2. Main section

The main instrument of this study is Bayes theorem:

$$P(S | H) = \frac{P(H | S)P(S)}{P(H)} = \frac{P(H | S)P(S)}{P(H | S)P(S) + P(H | \bar{S})P(\bar{S})}.$$

Let H denote the event of a certain type hallucination in the patient: auditory, visual, combined auditory and visual or tactile. Let S denote the event of a positive schizophrenia diagnosis.

In the context of this study, the presented probabilities have the following definition:

$P(S) = 0.4\%$ — the lifetime probability of diagnosing schizophrenia in an individual [3].

$P(\bar{S}) = 99.6\%$ — probability of not having a diagnosis of “schizophrenia: throughout life.

$P(H | S)$ — the probability of observing hallucinations of a certain type in a patient, provided that schizophrenia was diagnosed.

$P(H | \bar{S})$ — the probability of observing hallucinations of a certain type in a patient, provided that schizophrenia was not diagnosed.

The aim of this study was to observe the changes of a posterior probability, that indicates the existence of schizophrenia, given that certain types of individual’s hallucinations were being observed, meaning the calculation of probability $P(S | H)$.

Based on the data analysis of medical papers published in PubMed and JAMA Psychiatry, the probabilities of hallucinations in an individual were obtained given the presence/absence of schizophrenia. After applying Bayes theorem the values presented in column $P(S | H)$ of Table 1 have been received.

Probability table

Table 1

<i>H</i> , hallucinations observed	$P(H S)$	$P(H \bar{S})$	$P(S H)$
Auditory	83.4% [4]	2.5% [5]	11.814%
Visual	57.2% [4]	3.8% [5]	5.700%
Auditory and visual	84.0% [6]	0.2% [5]	62.780%
	$P(H S)$	$P(H)$	$P(S H)$
Tactile	27.0% [4]	3.1% [7]	3.483%

3. Conclusions

As a result of the meta-analysis of medical papers, the probabilities necessary for the application of Bayes theorem were determined, the required posterior probabilities clearly demonstrating the degree of influence of the observed signs on the possible final diagnosis were calculated.

Having calculated the posterior probabilities in the presence of auditory, visual, combined auditory and visual as well as tactile hallucinations, it could be concluded that the lowest comparative weight in the diagnostics had tactile hallucinations. At the same time, auditory and visual hallucinations had the greatest combined weight.

This study illustrates the possibility of applying the Bayes theorem as a tool for diagnosing mental illness. In addition, the above-described theoretical apparatus could be used to establish priority in the admission of patients in order to provide medical assistance promptly in a potentially dangerous situation.

References

1. WHO. Mental health: strengthening our response. — <http://www.who.int/mediacentre/factsheets/fs220/en/>. — [Online; accessed 2018-03-15].
2. van Os J., Kapur S. Schizophrenia // *Lancet*. — 2009. — Aug. — Vol. 374, no. 9690. — P. 635–645.
3. Bhugra D. The global prevalence of schizophrenia // *PLoS Med*. — 2005. — May. — Vol. 2, no. 5. — P. e151; quiz e175.

4. Thomas P., Mathur P., Gottesman I. I., Nagpal R., Nimgaonkar V. L., Deshpande S. N. Correlates of hallucinations in schizophrenia: A cross-cultural evaluation // Schizophrenia research. — 2007. — May. — Vol. 92, no. 1-3. — P. 41–49.
5. McGrath J., Saha S., Al-Hamzawi A., Alonso J., Bromet E., Bruffaerts R., Caldas-de Almeida J., Chiu W., de Jonge P. et al. Psychotic experiences in the general population: A cross-national analysis based on 31 261 respondents from 18 countries // JAMA Psychiatry. — 2015. — Vol. 72, no. 7. — P. 697–705. — Access mode: <http://dx.doi.org/10.1001/jamapsychiatry.2015.0575>.
6. Waters F., Collerton D., ffytche D. H., Jardri R., Pins D., Dudley R., Blom J. D., Mosimann U. P., Eperjesi F. et al. Visual Hallucinations in the Psychosis Spectrum and Comparative Information From Neurodegenerative Disorders and Eye Disease // Schizophrenia Bulletin. — 2014. — Vol. 40, no. Suppl_4. — P. S233–S245. — Access mode: <http://dx.doi.org/10.1093/schbul/sbu036>.
7. Ohayon M. M. Prevalence of hallucinations and their pathological associations in the general population // Psychiatry Research. — 2000. — Dec. — Vol. 97, no. 2-3. — P. 153–164.

УДК 519.248:616.895.8

Вычисление апостериорных вероятностей в психопатологии на основе метаанализа медицинских публикаций

А. Н. Козловский, В. В. Жуков

*Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: arcozlovskij@yandex.ru, zhukov_vv@rudn.university

Психическое здоровье человека, являясь достаточно сложным механизмом, требует квалифицированного обращения со стороны психиатров. Однако, в условиях ограниченного времени на прием и большого количества поступающих пациентов, оказывать каждому в равной степени квалифицированную помощь может быть затруднительно.

В данной работе рассматривается применение теоремы Байеса в диагностическом аппарате психопатологии, рассчитываются апостериорные вероятности диагностирования шизофрении при условии наблюдения у пациента одного или нескольких видов галлюцинаций, проводится оценка степени влияния признаков. Исходные статистические данные получены на основе метаанализа медицинских публикаций. Наибольшая апостериорная вероятность шизофрении наблюдается при условии наблюдения совмещенных слуховых и зрительных галлюцинаций (62.78%); наименьшая — при наблюдении тактильных галлюцинаций (3.48%).

Полученные результаты могут быть использованы для диагностирования заболевания или установки приоритета приема пациентов, что может существенно снизить нагрузку, возлагаемую на психиатров.

Ключевые слова: галлюцинации, шизофрения, психопатология, теорема Байеса, метаанализ.

УДК 004.4

Автоматизация оценивания вопросов открытого типа

Е. Ю. Коняев

*Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: 1032143099@rudn.university

Работа посвящена проблемам автоматизации проверки знаний обучающихся, в частности вопросам автоматизации проверки результатов тестов открытой формы.

Ключевые слова: автоматизация процесса обучения, автоматизация контроля знаний, система тестирования, открытые тесты.

1. Введение

В последние годы наблюдается тенденция автоматизации процесса обучения. Немаловажную роль в автоматизации процесса обучения играет автоматизация контроля знаний обучающихся. Автоматизация контроля знаний, в свою очередь, в большинстве случаев достигается за счёт организации тестирования. Чаще всего можно встретить тесты с выбором вариантов ответа, однако за счёт предоставления готовых вариантов ответа в них немалая вероятность угадывания. Так, в работе [1] было проведено следующее исследование: сначала обучающимся давались вопросы открытого типа, затем те же самые вопросы с выбором вариантов ответов. Эксперимент показал, что тесты с выбором ответов дают более высокую оценку за тест. Средняя разница в оценке по пятибалльной шкале по 3 группам составила 0,86 балла. Проведённый эксперимент говорит о том, что открытые тесты более объективно оценивают уровень знаний обучающихся.

Одна из причин, по которой открытые тесты используются реже, является сложность автоматизации оценивания ответов тестируемых. Как следствие, рассмотренные нами системы тестирования (Let's test [2], StartExam [3], INDIGO [4] и Moodle [5]) обладают ограниченным функционалом для работы с открытыми тестами. Таким образом, вопрос о создании системы тестирования для работы со всеми видами открытых тестов остаётся актуальным, чему и будет посвящена эта работа.

2. Виды открытых тестов и способы их оценивания

Прежде всего следует различать открытые тесты с числовыми и текстовыми ответами.

Если ответ представим в виде числа, то он будет считаться правильным, если введённое тестируемым число совпадает с эталонным ответом. Также, в зависимости от вопроса, может задаваться погрешность ответа, тогда для оценки истинности ответа проверяется его принадлежность промежутку (1)

$$E - \varepsilon \leq A \leq E + \varepsilon, \quad (1)$$

где A — ответ тестируемого, E — эталонный ответ, а ε — заданная погрешность.

Текстовые ответы, вообще говоря, могут быть двух типов: свободно конструируемые (когда тестируемый может дать ответ в свободной форме) и задания на дополнение (предполагают определённую форму ответа). Реализация автоматической проверки тестов первого типа довольно сложна, а в общем случае на данный момент невозможна. Задания второго типа поддаются автоматизации. Оценка истинности ответа тестируемого в данном случае зависит от метода оценки (дифференцированный или недифференцированный).

В случае недифференцированного метода оценки за задание ставится либо максимальный балл, либо 0. В этом случае данный тестируемый ответ сравнивается с эталонным на полное совпадение.

В случае дифференцированного метода оценки за задание ставится балл из диапазона [0; максимальный балл] в зависимости от совпадения данного тестируемого ответа с эталоном (полного, частичного или несовпадения). Совпадение может быть формальным и смысловым. В случае смыслового совпадения преподавателю самому необходимо создать различные эталоны и указать для них оценку из диапазона [0; максимальный балл]. Если же требуется оценить формальное совпадение данного тестируемого ответа с конкретным эталоном, то подразумевается, что ответ является множеством (порядок образующих ответ элементов неважен) или списком (порядок важен).

Пусть ответ на задание представляет из себя множество. Тогда оценку g_1 за задание можно рассчитывать по формуле (2):

$$g_1 = \frac{K_A}{L_E + K}, \quad (2)$$

где K_A — количество элементов из ответа тестируемого, входящих в эталонный ответ, L_E — количество элементов в эталонном ответе, K — количество элементов из ответа тестируемого, не входящих в эталонный ответ. Таким образом, если ответ тестируемого полностью совпадает с эталонным и не содержит лишнего, тестируемому ставится максимальный балл, если вообще не совпадает, то 0. Чем меньше элементов из эталона в ответе и чем больше в ответе лишних элементов, тем меньшую оценку получает тестируемый.

Пусть ответ на задание представляет из себя список, т.е. важен порядок элементов в ответе. Тогда оценку g_2 можно определять исходя из количества K_i перестановок, необходимых для приведения ответа тестируемого к эталонному ответу по следующей формуле (3):

$$g_2 = 1 - \frac{K_i}{K_n}, \quad (3)$$

где $K_n = \frac{n(n-1)}{2}$ — максимально возможное число перестановок для списка длиной n . Здесь предполагается, что ответ тестируемого содержит все и только все элементы из эталонного ответа и требуется оценить лишь их порядок. В противном случае нужно из исходных двух списков образовать другие два списка одинаковой длины, образованных одинаковыми элементами. Если исходные списки одинаковой длины, но часть образующих элементов одного списка не образует другой, то все отличные элементы и в одном и в другом списке заменяются на одинаковый произвольный элемент. По образованным спискам получается оценка g_2 , а по исходным g_1 . Итоговая оценка определяется как $g_1 * g_2$. Если списки разной длины, то в них также все отличные элементы заменяются одинаковым произвольным элементом и меньший список доводится до длины большего путём добавления того же произвольного элемента справа. Исходная оценка определяется аналогичным предыдущему случаю образом.

3. Заключение

Рассмотренные в данной работе методы оценивания заданий открытого типа были реализованы на РНР. Это позволило положить начало разработке системы тестирования в виде веб-приложения.

Литература

1. Карпова И. П. Сравнение открытых и выборочных тестов // Открытое образование. — 2010. — № 3. — С. 32–38.
2. Система тестирования Let's test [Электронный ресурс] / Режим доступа: <https://letstest.ru/>.
3. Система тестирования StartExam [Электронный ресурс] / Режим доступа: <http://www.opentest.ru/>.
4. Система тестирования INDIGO [Электронный ресурс] / Режим доступа: <http://www.indigotech.ru/>.
5. Open-source learning platform Moodle [Electronic resource] / Mode of access: <https://moodle.org/>.

UDC 004.4

Automated Scoring of Short-Answer Questions

E. Yu. Konyaev

*Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: 1032143099@rudn.university

This paper presents methods for automated scoring of short answer questions.

Key words and phrases: automation of the learning process, automated scoring, LMS, short-answer questions.

УДК 004.021

Выявление источников заимствования для документа с использованием моделей дистрибутивной семантики

А. Д. Крескин

*Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: leshakreskin@gmail.com

В статье предложен метод выявления источников заимствований для документа, использующий модель дистрибутивной семантики для формирования множества запросов к поисковой машине. Рассмотрены основные способы выявления заимствований и их источников. Показано, как выделить из документа запросы для поиска источников, используя векторное пространство, построенное на большом корпусе текстов при помощи инструмента Word2Vec. Приведены результаты работы метода.

Ключевые слова: дистрибутивная семантика, выявление заимствований, source retrieval, Word2Vec, антиплагиат.

1. Введение

С того момента, как люди стали создавать свои собственные творения другие люди пытались их копировать. Такие действия, умышленно совершаемые физическими лицами незаконное использование или распоряжение охраняемыми результатами чужого творческого труда, которое сопровождается доведением до других лиц ложных сведений о себе как о действительных авторах называются плагиатом.

Существуют несколько основных способов выявления заимствований. Одним из первых и самых простых был BoW (Bag of Words). Он основывался на представлении документа как некоего множества слов, а заимствования искал с помощью сравнения множеств для нескольких текстов. Такой подход хорошо работал в случае прямых и перефразированных заимствований, однако стоило лишь заменить слова синонимами, как метод переставал работать.

На помощь пришел семантический анализ текстов. В таком подходе основное внимание уделяется не столько самим словам, сколько их смыслу. Таким образом становится возможным искать синонимические, перефразированные и даже переводные заимствования.

В этой статье я опишу один из вариантов системы, выявляющей источники, из которых были произведены заимствования в документ и использующей инструменты семантического анализа.

2. Основная часть

2.1. Используемые инструменты

Для создания программы, был выбран язык программирования Python 2.7. Для представления слов в виде векторного пространства и для дальнейшего взаимодействия с ним – технология Word2Vec, с использованием библиотек gensim и numpy. В качестве поисковой машины была выбрана система ChatNoir, разработанная в Weimar Bauhaus University. Для тестирования был использован их же сервис TIRA.IO.

2.2. Принцип работы алгоритма

Повторю, что алгоритм предназначен для выявления источников заимствования в подозрительном тексте, поэтому его основная задача – получить высокую полноту результатов, при, возможно, не очень высокой точности.

2.2.1. Подготовка и загрузка W2V модели

Для работы алгоритма необходима качественная настроенная векторная модель, подготовка которой на домашнем компьютере отняла бы слишком много времени. Поэтому для работы алгоритма была выбрана модель, обученная разработчиком алгоритма Томашем Миколовым на корпусе текстов новостей компании Google.

Загрузка и взаимодействие с моделью выполняется при помощи библиотеки `gensim`

2.2.2. Нормализация и разбиение проверяемого текста на субпредложения

Нормализация происходит следующим образом. Текст очищается от всех знаков препинания, переводится в нижний регистр и разбивается на субпредложения длины `length` с пересечением `inter`. По сути `length` — длина окна, которое проходит по тексту, а `inter` — его шаг. После разбиения текста на субпредложения каждое из них фильтруется от повторов слов внутри себя. На этом токенизация заканчивается.

2.2.3. Выделение запроса для поисковой машины

Действия, описанные в этом параграфе, выполняются для каждого субпредложения полученного на предыдущем этапе.

Для выделения запросов алгоритм требует загруженную модель W2V. Сначала класс сопоставляет каждому слову субпредложения вектор из пространства векторов модели. Таким образом для данного субпредложения мы получаем некоторый кластер векторов. Далее по векторным представлениям слов вычисляется средний вектор — центр кластера.

Если рассматривать это действие на семантическом уровне, то таким образом вычисляется некое абстрактное понятие, к которому тяготеет весь кластер.

Следующим шагом алгоритм вычисляет векторное расстояние от каждого слова до центра кластера. Слова, расположенные недалеко от центра кластера, не представляют особого интереса, тогда как далеко отстоящие от центра слова являются характеризующими для данного кластера. Таким образом поисковый запрос составляется из слов, наиболее отстоящих от центра кластера.

2.2.4. Помещение запроса в поисковую машину и загрузка результатов

Не вижу смысла приводить здесь полный код работы с поисковой машиной. Скажу лишь, что в качестве таковой используется система ChatNoir. Для отправки запросов в систему, а также для получения результатов используются JSON-объекты.

2.3. Результаты тестирования алгоритма

Алгоритм показал не очень хорошие результаты на тестовых выборках. Ему явно необходима отладка и доработка. Однако даже на первых запусках были получены следующие результаты (табл. 1).

Как и предполагалось, показатель Recall намного превышает показатель Precision. Однако, показатель полноты тоже достаточно низок. Предположительно это связано с недостаточной фильтрацией запросов.

3. Заключение

В этой работе было показано, как с помощью достаточно простых средств можно построить систему выявления источников заимствований для текстов. Несомненно

эта система не сравнится с промышленными гигантами, как «Антиплагиат», однако при должной доработке может показать хорошие результаты.

Таблица 1

Результаты работы алгоритма

Settings	Precision	Recall	F-measure	Queries quantity
ql=6, len=50, inter=25	0.015	0.355	0.029	651
ql=5, len=80, inter=45	0.02	0.25	0.036	462

Литература

1. Meuschke N. et al. Analyzing Semantic Concept Patterns to Detect Academic Plagiarism // Proceedings of the 6th International Workshop on Mining Scientific Publications. — ACM, 2017. — Pp. 46–53.
2. Tian Z. et al. Software plagiarism detection with birthmarks based on dynamic key instruction sequences // IEEE Transactions on Software Engineering. — 2015. — Vol. 41, no. 12. — Pp. 1217–1235.
3. Sochenkov I. et al. Exactus Like: Plagiarism detection in scientific texts!// European Conference on Information Retrieval. — Springer, Cham, 2016. — Pp. 837–840.
4. Sochenkov I. V., Maluleka R. Query Formulation for Source Retrieval based on Named Entities and N-grams Extraction // Искусственный интеллект и принятие решений. — 2016. — №. 3. — С. 44–47.
5. Potthast M., Gollub T. Rangel F. Rosso P. Plagiarism Detection, Author Identification, and Author Profiling. — 2014.
6. Mikolov T., Sutskever I., Chen K., Corrado G. S., Dean J. Distributed representations of words and phrases and their compositionality // Advances in neural information processing systems. — 2013. — Pp. 3111–3119.

UDC 004.021

Source Retrieval for a Document Using Models of Distributive Semantics

A. D. Kreskin

*Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: leshakreskin@gmail.com

This paper is about method for identifying sources of plagiarism for a document, using a model of distributive semantics to form a set of queries to a search engine. The main ways of revealing plagiarisms and their sources are considered. It shows how to select queries from the document to search for sources using a vector space built on a large body of texts using the Word2Vec tool. The results of method's work are presented.

Key words and phrases: source retrieval, distributive semantic, Word2Vec, plagiarism detection.

УДК 681.3.016

Проектирование многомерных информационных систем с использованием методологии «Data Vault»

Е. А. Кузнецов*, М. Б. Фомин†

* Департамент цифровых решений

Лаборатория новых информационных технологий (ЛАНИТ)

Мурманский пр-д, д. 14, Россия, 129075

† Кафедра информационных технологий

Российский университет дружбы народов

ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198

Email: kuznetsovea@lanit.ru, fomin_mb@rudn.university

Рассмотрен способ проектирования информационно-аналитических систем с использованием методологии «Data vault». Такой подход позволяет строить хранилище данных информационной системы на основе метамодели, которая семантически связана с предметной областью системы, легко перестраивается в случае изменений в бизнес-модели предметной области и позволяет формировать витрины данных на принципах OLAP.

Ключевые слова: хранилище данных, многомерная модель данных, витрина данных, OLAP, data vault.

1. Введение

Удешевление высокопроизводительных вычислительных комплексов сделало их доступными для предприятий среднего бизнеса, функционирование которых связано с выполнением большого объема разнотипных операций. Возникла потребность в недорогих и простых в эксплуатации информационно-аналитических системах (ИАС), обеспечивающих потребности анализа деятельности предприятий. Такие ИАС должны удовлетворять следующим требованиям:

- система должна обрабатывать данные, требуемые в процессе принятия решений по управлению предприятием;
- система должна позволять описывать и производить расчет ключевых показателей эффективности по видам деятельности предприятия;
- структура метаданных хранилища данных должна соответствовать бизнес-процессам предприятия;
- должна быть возможность оперативного внесения в систему изменений при изменениях в деятельности предприятия или в случае внесения изменений в методику анализа деятельности.

2. Сценарий работы информационной системы

Деятельность информационной системы, удовлетворяющей сформулированным требованиям, может быть представлена в виде конвейера движения данных со сценарием работы, схема которого изображена на рис. 1.

ИАС состоит из следующих подсистем: подсистема сбора данных (ПСД), хранилище данных (ХД), аналитическая подсистема. На первом этапе движения данных ПСД взаимодействует со смежными информационными системами и формирует корректный контент оперативного склада данных (ОСД) – временной области хранения, в которой данные существуют до следующего шага – загрузки в первичный слой данных ХД. Загрузка данных ОСД в ХД производится путем их сопоставления с актуальной моделью данных ХД.

Управление моделью данных должно осуществляться на уровне метамодели. Метаданные должны быть описаны таким образом, чтобы имелась возможность задания на их основе способа расчета показателей эффективности, используемых в процессе анализа деятельности предприятия.



Рис. 1. Сценарий работы информационно-аналитической системы

На основе первичного слоя данных осуществляется расчет показателей, которые размещаются в специальные структуры хранения данных – тематические витрины данных аналитической подсистемы. В качестве модели представления этих данных удобно использовать многомерную модель, основными элементами которой являются факты и измерения. Такое представление дает возможность использовать в качестве механизмов предоставления данных для конечного пользователя концепцию OLAP.

3. Описание метамодели данных с использованием методологии data vault

Одним из способов построения хранилища данных является методология data vault. Её использование позволяет динамически расширять модель данных ХД без возникновения сложной задачи модификации других подсистем ИАС. Управление моделью данных должно осуществляться на уровне метамодели. Основные объекты метамодели: бизнес ключ (в терминологии data vault – «hub»), транзакция бизнес ключа (в терминологии data vault – «link») и история бизнес ключа (в терминологии data vault – «sat»). Бизнес ключ – свойство объекта, однозначно идентифицирующее его в рамках предметной области. История бизнес ключа – история изменений свойств объекта, функционально зависящих от этого бизнес ключа. В рамках многомерной модели данных в аналитической подсистеме бизнес ключ может выступать в качестве измерения. Актуальность атрибутов измерения поддерживается на основе истории бизнес ключа. Транзакция по бизнес ключам – описание события, произошедшего между объектами, которые идентифицируются этими бизнес ключами. История транзакций по бизнес ключам позволяет производить расчет показателей в многомерных моделях данных, используемых в тематических витринах. На рис. 2 представлена структура метамодели ХД в виде диаграммы в нотации E/R+Merise.

Для того чтобы в процессе модификации ИАС не была потеряна информация об имеющихся в метамодели ассоциативных связях по типу связи (агрегация, композиция или рекурсия) и по арности ($1 : 1$, $1 : N$, $M : N$), эта информация должна сохраняться транзакциях.

4. Заключение

В работе был рассмотрен способ проектирования информационно-аналитических системах с использованием методологии «Data vault». Такой подход позволяет строить хранилище данных системы на основе метамодели, которая семантически связана с предметной областью системы, легко перестраивается в случае изменений в бизнес-модели предметной области и позволяет формировать многомерные витрины данных.

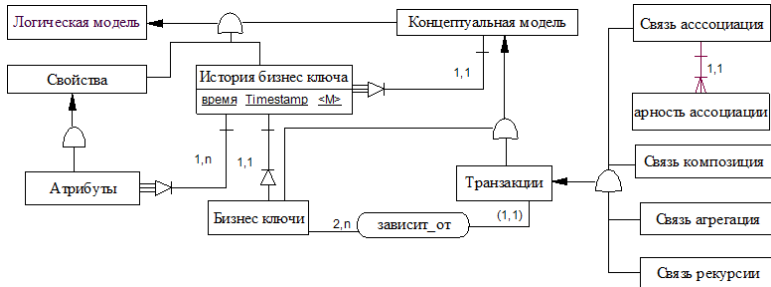


Рис. 2. Структура метамодели хранилища данных

Литература

1. Linstedt D., Olschimke M. Building a Scalable Data Warehouse with Data Vault 2.0, Elsevier Science, 2015.

UDC 681.3.016

Designing Multidimensional Information Systems Using the Data Vault Methodology

E. A. Kuznetsov*, M. B. Fomin†

* Department of digital solutions
 Laboratory of New Information Technologies (LANIT)
 14 Murmanskii proezd, Moscow, 129075, Russian Federation

† Department of Information Technology
 Peoples' Friendship University of Russia (RUDN University)
 6 Miklukho-Maklaeva str., Moscow, 117198, Russian Federation

Email: kuznetsovea@lanit.ru, fomin_mb@rudn.university

The method of designing information and analytical systems using the data vault methodology is considered. This approach allows building an information system data warehouse based on the basis of meta-model, which is semantically related to the information system domain. The structure of the information system is easily rebuilt in case of changes in the business model of system domain and allows you to form data marts on the principles of OLAP.

Key words and phrases: data warehouse, multidimensional data model, data mart, OLAP, data vault.

УДК 519.688

Построение системы проведения мета-анализа медицинских исследований для получения ответа на клинический вопрос (на примере печеночной недостаточности)

А. В. Лукин*†

** Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклуко-Маклая, д. 6, Москва, Россия, 117198*

*† ООО «Технологии системного анализа»
проспект 60-Летия Октября, д. 9, Москва, Россия, 117312*

Email: antonvlukin@gmail.com

В настоящей статье рассматривается задача построения системы проведения мета-анализа данных выделенных из медицинских публикаций для получения ответа на клинический вопрос. Для проведения мета-анализа необходимо большое количество гомогенных данных, выделяемых из медицинских статей. Процесс выделения данных начинается с разметки экспертами небольшого числа статей с целью выделения структуры значимых медицинских данных. Затем опираясь на выделенную структуру составляется модель извлечения данных, содержащая в себе набор ключевых слов, правила отбора данных, синонимические словари. На основе эффективной модели извлечения данных из основного корпуса релевантных публикаций могут быть выделены данные в количестве необходимом для проведения мета-анализа.

Ключевые слова: компьютерные науки, информационные технологии.

1. Введение

В области медицины, каждый день публикуются новые статьи, материалы, отчеты о проведенных исследованиях. Вопросы подбора лечения непосредственно связаны с клиническими проявлениями болезни, которые определяют подбор лечебных препаратов. К примеру, больных с печеночной недостаточностью, можно разделить на группы, определяющиеся проявлением релевантных признаков. На этой основе можно подбирать оптимальную терапию. Но для этого нужны аналогичные по проявлениям группы с завершенными результатами лечения. Такие группы можно формировать по литературным данным с помощью мета-анализа [1, 2].

Актуальность работы заключается в необходимости автоматизации получения ответа на клинический вопрос, возникающего в процессе лечения пациента, в упрощении процесса подбора максимально эффективной стратегии лечения для каждого уникального пациента.

Объектом исследования является механизм проведения мета-анализов в области доказательной медицины.

Цель исследования — разработка интеллектуальных методов поддержки мета-анализов.

Задачи исследования — описание процесса автоматического подбора публикаций релевантных клиническому вопросу, разработка архитектуры системы интеллектуальной поддержки проведения мета-анализов.

2. Основная часть

Построение системы, предназначенной для проведения мета-анализов состоит из двух этапов:

1. Подготовка обучающей выборки.
2. Построение модели для отбора публикаций отвечающих на клинический вопрос [3].

В качестве исходных данных были использованы вручную отобранные статьи о печеночной недостаточности. Статьи были размечены экспертами (врачами) с целью выделения структуры значимых медицинских данных.

Для проведения последующего мета-анализа необходимо чтобы данные выделенные из статей отвечали высокой степени доказательности. Для этого при первичном поиске в медицинских базах данных указывается тип проводимого исследование - Randomized controlled trial (RCT) или Double-blind randomized controlled trial (DB-RCT). Так же на этапе выделения положительных кандидатов из неразмеченных работ, статья дополнительно проверяется классификатором [4].

Архитектура системы построения мета-анализов:

ЭТАП 1. Врач составляет клинический вопрос, формулирует клиническую ситуацию на естественном языке в свободной форме и выбирает тип вопроса. Вопрос в автоматическом режиме переводится на английский язык, затем преобразуется в формат PICO, и выводится на экран для корректировки. При этом система использует медицинский тезаурус «MeSH» отображает подсказки на русском и английском языках. Пользователь корректирует поисковые поля, и система транслирует термины из формата PICO в поисковый запрос.

ЭТАП 2. Производится поиск информации в электронных библиотеках медицинской направленности (Trip Database, Cochrane Library, МЭБ, PubMed). Поисковый механизм в автоматическом режиме направляет поисковый запрос к информационным ресурсам. Результаты поиска выводятся на экран. Задача данного этапа - первичный отбор публикаций релевантных поставленному клиническому вопросу. Отобранные публикации будут сохранены в системе.

ЭТАП 3. Производится оценка степени семантической релевантности найденных документов.

Критический анализ отобранных публикаций проводится поэтапно:

- оценка релевантности путем сопоставления запроса в формате PICO, составленного пользователем, с материалами и методами анализируемой публикации;
- определение дизайна исследования в соответствии с алгоритмом, включающим последовательность контрольных вопросов;
- оценка качества исследования с помощью контрольных вопросов и балльной оценки;
- интерпретация количественных результатов с возможностью расчета дополнительных показателей по данным исследования.

3. Заключение

В заключении хотелось бы отметить что исследование на данный момент не завершено, с целью более точного и полного выделения значимой медицинской информации из научных текстов производится доработка правил и модели полуавтоматической разметки. Так же для формирования более подробной обучающей выборки от экспертов ожидается более 20 дополнительно размеченных научных статей.

На данный момент экспертами размечены 54 научные статьи. На основе выделенной структуры разметки разработаны 14 правил выделения информации, составлено 3 пары словарей с положительными и негативными ключевыми словами, построена модель отбора релевантных публикаций.

Благодарности

Работа частично поддержана грантом РФФИ № 16-29-07354.

Литература

1. Кобринский Б. А., Молодченков А. И., Благосклонов Н. А., Лукин А. В. Применение методов метаанализа в диагностике и лечении пациентов с печеночной

- недостаточностью. Программные продукты и системы. Тверь, Россия, 2017. doi:10.15827/0236-235X.030.4.745-753.
2. Dans A. L., Dans L. F., Silvestre M. A. A. *Painless evidence-based medicine*. Chichester. Wiley, 2008.
 3. Boyko A. A., Kaidina A. M., Kim Y. C., Lupatov A. Yu., Panov A. I., Suvorov R. E., Shvets A. V. *A Framework for Automated Meta-Analysis: Dendritic Cell Therapy Case Study*. Institute of Biomedical Chemistry of the Russian Academy of Medical Sciences. Moscow, Russia, 2016. doi:10.1109/IS.2016.7737416.
 4. Guyatt G., Rennie D., Meade M. O., Cook D. J. *Users' guide to the medical literature. A manual for evidence-based clinical practice – 2nd ed.* McGraw-Hill, 2008.

UDC 519.688

Development of the System for Conducting a Meta-Analysis of Medical Research to Obtain a Response to a Clinical Question (on the Example of Liver Failure)

A. V. Lukin*[†]

** Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

*† Technologies for Systems Analysis Ltd.
60-Letiya Oktyabrya prospectus, Moscow, 117312, Russian Federation*

Email: antonvlukin@gmail.com

In this article, we consider the task of constructing a system for conducting a meta-analysis of data extracted from medical publications to obtain a response to a clinical question. To conduct a meta-analysis, a large amount of homogeneous data is required from medical articles. The process of data extraction begins with the marking by experts of a small number of articles in order to highlight the structure of significant medical data. Then, based on the selected structure, a data extraction model is compiled, containing a set of keywords, data selection rules, and synonymous dictionaries. Based on an effective model of extracting data from the main body of relevant publications, data can be extracted in the amount necessary for conducting a meta-analysis.

Key words and phrases: computer science, information technologies.

УДК 004

Методы синтеза текстов на естественном языке в заданном стиле

Е. А. Матвеева

*Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: eam4tveeva@gmail.com

В данной работе исследуется задача изменения стиля текстов на естественном языке. Целью исследования является разработка методики «перевода» текстов из одного стиля или жанра в другой. В работе рассматриваются теоретические вопросы генерации текстов в определенном стиле или жанре и методы замены стиля в заданном тексте, проведен обзор существующих решений и идей по решению этой задачи. Предложен метод для реализации идеи автоматического замены стиля текста с использованием векторных представлений слов. В практической части данного исследования описан эксперимент по переводу стиля текста на основе разработанного метода с использованием корпусов Википедии на русском языке и подборки русскоязычных новостей.

Ключевые слова: обработка естественного языка, синтез текста, word2vec.

1. Введение

Задачи синтеза текстов на естественных языках не теряют своей актуальности и напротив даже, растут в популярности с каждым днем. Одной из наиболее распространенных задач этой области является задача генерации не просто грамотного и имеющего смысл текста, но текста, написанного в каком-то определенном стиле или жанре. В данной работе задача генерации текста в стиле была сведена к более простой задаче «перевода» текста из одного стиля в другой. Целью данного исследования является разработка методики «перевода» текстов из одного стиля или жанра в другой. Для достижения поставленной цели было необходимо решить следующие теоретические и практические задачи:

1. Рассмотреть и описать уже существующие решения и идеи для решения поставленной задачи, определить их сильные и слабые стороны;
2. Разработать алгоритм решения задачи, выбрать инструменты для использования в реализации алгоритма;
3. Проверить работу данного алгоритма и оценить его эффективность.

Практическая значимость работы заключается в том, что предложенный метод может быть развит далее для решения более комплексных задач, а полученные на данном этапе результаты могут быть использованы, например, в качестве контента-наполнителя. После исследования возможных способов обработки и генерации текстов на естественных языках и особенности каждого из них было решено использовать группу алгоритмов для получения векторных представлений слов word2vec. Метод, разработанный для решения задачи «перевода», базируется на использовании векторных представлений текст-корпусов для дальнейшего поиска зависимостей между представлениями подобных слов в разных корпусах. В представленной работе описаны теоретические идеи метода, созданного для решения проблемы, его разработка, а также практические эксперименты, получение некоторых конкретных результатов и их оценка.

2. Проблема синтеза текстов в заданном стиле

Задача генерации текстов в стиле стоит перед исследователями не первый год. Под стилем имеются в виду некоторые ожидания или предпочтения конкретного пользователя о способе отображения контента и/или жанр документа, то есть не просто контент, но контент, опирающийся на определенную структуру текста и языка.

Сегодня можно найти множество применений рекуррентных нейронных сетей для генерации текста в стиле, однако, сгенерированные посимвольно, тексты получаются бессмысленными [1]. Что хочется видеть в искусственных текстах — это связь между синтаксисом и семантикой языка. Именно поэтому в данной работе была выделена задача “смены” стиля текста. Это первый шаг к генерации жанровых текстов, но на этом этапе можно ожидать более или менее осмысленные результаты, ведь за основу будет взят текст, написанный человеком, а машина будет его лишь модифицировать.

К сожалению, несмотря на многочисленные прорывы в области обработки естественных языков, очень мало внимания было уделено непосредственно пониманию стиля текста. В недалеком прошлом серьезный прогресс был достигнут в области изменения стиля изображения — одним из примеров является приложение Prisma. Однако можно найти очень мало работ, преследующих ту же цель относительно текста.

Большинство новых идей основывается на использовании модели для машинного перевода seq2seq. Но главной проблемой остается необходимость найти «параллельные» выборки текстов для обучения, одинаковых по смыслу, но различающихся стилистически. Таких примеров мало, и, так как seq2seq требует именно параллельные данные, эта проблема делает задачу довольно сложной. Впрочем, комбинируя различные методы и подходы, исследователям из Стэнфорда удалось получить некоторые результаты именно с помощью seq2seq [2].

В 2017 году группа анонимных исследователей опубликовали идею фреймворка типа «encoder-decoder» [3] для переноса стиля с одного текста на другой и получили неплохие результаты, правда, для китайского языка, что заставляет делать некоторые оговорки в связи с уникальностью китайского языка, его лексики и грамматики.

Еще одна интересная работа на иную, но очень близкую тему, была опубликована в 2016 году в Эдинбурге [4] — здесь рассматривается задача перевода с одного естественного языка на другой, сохраняя при этом стиль повествования.

Самая последняя обнаруженная работа на эту тему предлагает похожий на описанный ниже подход: берутся формальные модели двух «стилей», сравниваются их векторные представления, и с помощью машинного перевода влияют на выходные данные, тем самым изменяя автоматические метрики [5].

Как видно, работ на данную тему на сегодняшний день немного, а очевидных результатов — тем более. В данной работе предлагается подход с использованием модели word2vec и искусственных нейронных сетей.

Word2vec [6] получает большой текстовый корпус в качестве входных данных и вычисляет векторное представление всех слов, основываясь на контекстной близости: слова, встречающиеся в исходном тексте рядом с одинаковыми словами (а значит, имеющие схожий смысл), в векторном пространстве будут располагаться недалеко друг от друга. В качестве инструмента будет использоваться язык python и библиотека gensim [6], в которой реализовано обучение моделей для word2vec.

При разработке метода необходимо было учитывать, что найти для обучения большую подборку одинаковых текстов, написанных в разных жанрах, представляется затруднительным, поэтому остается вариант обучаться на так называемых стоп-словах. Предложенный метод базируется на определении широко распространенных слов в текстах разных стилей, а затем сопоставлении их в векторных пространствах каждого текста. После получения набора одинаковых или очень близких по смыслу слов в векторных их представлениях, остается только определить, как они связаны между собой. Иными словами, задача сводится к нахождению некоторого оператора A , преобразующего вектор V_1 (представление слова в тексте 1) в вектор V_2 (представление слова в тексте 2). Необходимо учитывать, что для каждой пары слов оператор может немного отличаться, соответственно, нужно найти не просто оператор A , но какое-то его усредненное значение. Это можно сделать, построив простейшую нейросеть, входными данными которой будет множество векторов V_1 , а выходными — V_2 .

Наконец, полученные данные можно будет использовать для преобразования текстов из одного стиля в другой: конкретные слова помогут нам определить

правило, по которому они зависят друг от друга, а затем на основании этого правила мы сможем преобразовывать и другие, не входящие в обучающую выборку, слова.

3. Заключение

Данная работа представляет собой теоретический материал по выбранной теме и разработанному для решения выбранной задачи методу, а также результаты экспериментов, проведенных с использованием предложенного метода.

В ходе исследования проблемы было выяснено, что внимания задаче в настоящее время уделено недостаточно, именно поэтому данное исследование является актуальным.

Для решения проблемы был предложен метод, основанный на использовании модели word2vec и векторных представлений текстов, а также искусственных нейронных сетей для выявления необходимых закономерностей и правил для последующего использования на практике.

Реализация метода позволяет провести эксперименты, в ходе которых можно получить некоторые результаты и сделать соответствующие выводы.

Литература

1. Andrej Karpathy [Электронный ресурс]: <http://karpathy.github.io/>
2. Jang S.W., Min J., Kwon M.. Writing Style Conversion using Neural Machine Translation — Stanford University, 2017.
3. Language Style Transfer from Non-Parallel Text with Arbitrary Styles, 2017.
4. Sennrich R., Haddow B., Birch A. Controlling politeness in neural machine translation via side constraints. — University of Edinburgh, 2016.
5. Niu X., Martindale M., Carpuat M. A Study of Style in Machine Translation: Controlling the Formality of Machine Translation Output. — University of Maryland, 2017.
6. Word2vec [Электронный ресурс]: <https://en.wikipedia.org/wiki/Word2vec>
7. Genism.word2vec [Электронный ресурс]: <https://radimrehurek.com/gensim/models/word2vec.html>

UDC 004

Approaches to Natural Language Generation with Style

E. A. Matveeva

*Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: eam4tveeva@gmail.com

This paper focuses on content generation with style and style transfer on text in natural language. The key challenge of this work is to develop a method of transferring sentences from one style (or genre) to another. The paper deals with the theoretical issues of generating text in a certain style or genre and methods for changing style of a given text and contains a review of existing solutions and ideas for solving the stated problem, the proposed method for implementing the idea of automatic style transfer using vector representation of words and experiments on style transfer based on the developed method using Russian Wikipedia and Russian-language news corpuses.

Key words and phrases: natural language processing, text generation, word2vec.

УДК 004.4

Модуль базы знаний в системе управления здоровьем людей

А. И. Молодченков*†

** Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

*† Отдел «Интеллектуальные технологии и системы»
Федеральный исследовательский центр «Информатика и управление» РАН
Пр-т 60-летия Октября, д. 9, Москва, Россия, 117312*

Email: aim@tesyan.ru

В настоящее время в медицине активно развивается направление 4П медицины. Оно направлено на выявление рисков возникновения заболеваний с целью выработки профилактических мероприятий по снижению этих рисков. В работе описан модуль базы знаний системы оценки рисков возникновения отклонений в здоровье людей и построения плана профилактических мероприятий для снижения этих рисков.

Ключевые слова: 4П медицина, базы знаний, профилактика, искусственный интеллект.

1. Введение

В настоящее время активно развивается направление в медицине, которое получило название 4П медицина. Это направление базируется на следующих принципах [1, 2]:

- предиктивности (предсказательности), позволяющей прогнозировать заболевания;
- превентивности (профилактики), работающей на опережение и позволяющей предотвращать появление заболеваний с помощью их профилактики, а также вакцин и препаратов;
- персонализации, основанной на индивидуальном подходе к каждому больному;
- партисипативности (участия, партнерства), основанной на широком сотрудничестве различных врачей-специалистов и пациентов, а также на превращении пациента из субъекта лечения в объект лечебного процесса.

Современные работы в этом направлении ведутся только с использованием биотехнологий, в частности генома человека. Однако для эффективной поддержки здоровья людей, кроме данных о геноме и постоянной сдачи анализов, можно использовать и другую информацию. Например, образ жизни людей, двигательную активность, социальные сети, отдельные показатели здоровья и др. Все эти данные можно получать с помощью современных технологий мобильной медицины, Интернета вещей, анализа текстов и т.п. Возникает вопрос использования этих данных для выявления рисков возникновения заболеваний у человека с учетом того, что методы машинного обучения не могут решить эту задачу в полном объеме.

В работе описана структура базы знаний системы поддержки здоровья людей и его управления, путем выработки плана профилактических мероприятий, на примере инсульта и инфаркта.

2. Модуль базы знаний

Для оценки рисков развития заболеваний в разрабатываемом комплексе разработан модуль базы знаний. База знаний содержит разделы базы знаний, объекты, свойства объектов, связи между объектами и представляет собой компьютерную модель неоднородной семантической сети [3–5].

Объекты — основные составляющие базы знаний. Ими могут быть различные объекты предметной области: утверждения, факты, результаты наблюдений, анализов, рекомендации.

Объекты в базе знаний сгруппированы в разделы. Количество разделов не ограничено. Объект может иметь свойства, набор которых является характеристикой этого объекта.

В базе знаний объекты, определяющие область знаний, разбиваются на две группы. В первую группу попадают исходные объекты, от которых зависит ход рассуждения (они называются признаками), а во вторую — объекты, которые формируют решения — целевые объекты, иначе их называют гипотезами. В базе знаний системы признаки делятся на две основные группы: рисковые события и факторы риска. В качестве гипотез выступают риски возникновения заболеваний. В то же время выделяются следующие типы факторов риска: Постоянные, Управляемые, Предрасполагающие к развитию заболевания, Лабораторные, Выявляемые инструментальными методами. В качестве постоянных факторов риска могут выступать пол, возраст и др., управляемых — курение, злоупотребление алкоголем, низкая физическая активность и др., предрасполагающие — диабет, инфаркт а анамнезе и др., Лабораторные — холестерин, выявляемые инструментальными методами — артериальное давление, ЭКГ и др.

Для того чтобы наличие одних событий в предметной области могло повлечь порождение либо исключение других, между событиями в базе знаний устанавливаются связи. Связь можно установить между любыми двумя объектами, хотя наиболее важны связи между объектами — признаками и целевыми объектами — гипотезами. Например, «Периодическое повышение давления» может привести «Повышенный риск возникновения Инфаркта». Таким же образом связываются и другие факторы риска с исходами.

Вывод гипотез, признаков и ввод информации в базу осуществляется с помощью решателя. Этот решатель имеет унифицированный программный интерфейс (API) для того, чтобы от пользователей эта информация могла собираться различными способами: с применением технологий мобильной медицины, Интернета вещей, Web-приложения и др.

3. Заключение

В работе описана структура модуля базы знаний системы анализа состояния здоровья человека и выработки плана профилактических мероприятий для стабилизации и поддержания состояния здоровья человека на высоком уровне. Основной задачей разработанного модуля является комплексная оценка факторов риска состояния здоровья пользователя системы и выявления рисков возникновения различных заболеваний и обяснения выдвигаемых гипотез для дальнейшего снижения этих рисков. Разработанный модуль может применяться как внутри одной большой разрабатываемой системы, так и отдельными приложениями в области здоровья, разрабатываемыми сторонними разработчиками.

Благодарности

Исследования проводились при финансовой поддержке Министерства образования и науки РФ. Уникальный идентификатор соглашения RFMEFI60417X0194

Литература

1. Hood L., Flores M. A personal view on systems medicine and the emergence of proactive P4 medicine: predictive, preventive, personalized and participatory // *New biotechnology*. — 2012. — Vol. 29, no. 6. — Pp. 613–624.

2. 4п-медицина — новое направление развития здравоохранения (Электронный ресурс), URL: <http://federalbook.ru/files/FSZ/soderghanie/Tom%2013/II/Gerasimenko.pdf> (Дата обращения 05.03.2018).
3. Осипов Г. С. Приобретение знаний интеллектуальными системами: Основы теории и технологии. — М.: Наука. Физматлит, 1997. — 112 с.
4. Осипов Г. С. Методы искусственного интеллекта. — 2-е издание. — М.: Физматлит, 2015, 296 с.
5. Макаров Д. А., Молодченков А. И. Создание систем приобретения знаний для построения медицинских экспертных систем на основе ядра программных инструментальных средств «MedExp» // Информационно-измерительные и управляющие системы. — 2009. — Т. 7, № 12. — С. 86–89.

UDC 004.4

Knowledge Base Module of the System for Human Health Management

A. I. Molodchenkov*[†]

** Department of Information Technologies*

*Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

*[†] Department "Intelligent Technologies and Systems"
Federal Research Center "Computer Science and Control" of RAS
117312, Moscow, pr. 60-letiya Oktyabrya, 9, Russia*

Email: aim@tesyan.ru

At present time the direction of P4 medicine is actively advancing in medicine. It is aimed at identifying risks of disease and developing preventive measures to reduce these risks. The knowledge base module of the system for assessing the risks of deviations in human health and the construction of a plan of preventive measures to reduce these risks is described in this paper.

Key words and phrases: P4 medicine, knowledge base, preventive measures, artificial intelligence.

УДК 004.8

Разрешение неоднозначности авторства публикаций на основе коллекции документов поисковой системы Exactus Expert

Д. С. Новикова

*Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: dsnovikova@inbox.ru

В настоящей работе представлен обзор подходов и методов разрешения неоднозначности авторства публикаций и сделаны выводы о перспективности их использования, а также выбрано направление будущих исследований в указанной области на основе анализа больших коллекций текстовых документов.

Ключевые слова: неоднозначность авторства, библиографические данные, автоматическое разрешение неоднозначности, метрики близости научных документов, коллекции текстовых документов.

1. Введение

Проблема разрешения неоднозначности авторства публикаций в последние годы находится в фокусе внимания многих ученых. Повышенный интерес к указанной проблематике обусловлен ее непосредственным влиянием на точность извлечения информации из электронных библиотек и вычисление наукометрических показателей. По мнению автора работы, одним из существенных ограничений существующих в этой области подходов является использование только реферативных баз данных. В результате этого снижается вариативность признаков для публикаций. В нашем исследовании мы планируем использовать полнотекстовую базу данных, что позволит нам ликвидировать указанное ограничение.

2. Существующие методы решения проблемы неоднозначности авторства

Методы в области разрешения неоднозначности авторства публикаций формально можно разбить на две группы: требующие и не требующие обучения.

Среди методов, требующих обучения с учителем, стоит отметить метод [1], использующий глубокую нейронную сеть, и метод [2], использующий экстремальное обучение.

Среди методов, использующих обучение без учителя, стоит отметить алгоритм, предложенный в работе [3]. Данный алгоритм основан на теории Демпстера-Шафера в комбинации с энтропией Шеннона. На первом шаге он опирается на наиболее значимые признаки, такие как аффилиция автора, соавторы, место публикации, заголовков публикации и т.д. На втором шаге полученные показатели близости объединяются с использованием теории Демпстера-Шафера и энтропии Шеннона. На основе полученной информации для каждого автора происходит вычисление функции доверия и правдоподобия. В результате получается матрица попарной корреляции, с использованием которой производится иерархическая кластеризация.

В работе [4] предложен быстрый алгоритм кластеризации, также опирающийся на обучение без учителя, согласно которому на первом этапе происходит формирование кластеров с использованием имен соавторов, после чего внутри кластеров происходит установление связей, которые имеются между авторами и публикациями, а также между самими публикациями. На основе установленных связей происходит объединение связанных публикаций в отдельные кластеры, после чего

кластеры укрепляются с учетом близости заголовков публикаций. В случае если заголовки не содержат одинаковой ключевой лексики, для увеличения кластеров используется также место публикации.

Имеется ряд работ, например [6], в которых помимо текстовой оставляющей учитывается также и контекстное окружение, которое характеризуется наличием взаимосвязей между атрибутами публикации. В этом случае контекстуальное окружение публикации представляется в виде графа. Подход, основанный на анализе графов, получил дальнейшее развитие в работах [4] и [5]. В работе [4] для кластеризации используется метод распространения близости (Affinity propagation). Его очевидное преимущество состоит в том, что он не требует того, чтобы предварительно было известно или задано число кластеров. В работе [5] использованы топологические особенности графа совместной встречаемости наряду с совместными признаками авторов.

3. Направление будущих исследований

В качестве метода кластеризации в ходе работы было решено использовать иерархическую кластеризацию, так как она хорошо зарекомендовала себя в рамках многих алгоритмов. В качестве признаков для кластеризации планируется использовать, во-первых, стандартные признаки, такие как имена соавторов и заголовков публикации, а во-вторых, дополнительную информацию, характеризующую публикацию, такую как ключевая лексика. На данный момент уже имеется инструмент извлечения ключевой лексики и поиск по коллекции похожих работ [7].

В связи с тем, что при работе с большой коллекцией документов существует проблема ограничения вычислительных средств по памяти, в первую очередь, планируется проведение «мягкой» кластеризации с использованием одного высокоинформативного признака. Ведущая роль на этом этапе будет отведена ключевой лексике документа, а также семантической близости документов.

Занижение индекса цитирования для русскоязычных авторов часто происходит в связи с тем, что их публикации могут быть написаны не только на русском, но и на английском языке. В данном случае речь идет о проблеме омонимии, когда публикации одного автора воспринимаются как работы нескольких разных ученых. Для учета англоязычных работ в рамках работы планируется производить транслитерацию имен. Так как имена, полученные в итоге, могут незначительно отличаться от оригинала, то сравнение с имеющимися в базе именами планируется производить с использованием одной из стандартных мер близости строк.

Также в дальнейшем планируется использование таких сторонних признаков, как место работы автора и e-mail адреса. Использование перечисленных признаков предназначено, в первую очередь, для тех публикаций, которые выполнены без соавторов и/или извлечение из них ключевой лексики не дало результатов.

4. Заключение

Изучение имеющихся в исследуемой области материалов показало, что все встретившиеся в работах алгоритмы были разработаны для реферативных баз данных, таких как DBLP, CiteSeerX, PubMed, BDBCComp и других, которые содержат только библиографическую информацию о публикации. Таким образом, в них отсутствуют полные тексты статей, что уменьшает вариативность признаков для проведения кластеризации.

В отличие от существующих в исследуемой области подходов наш метод предполагает повышение точности разрешения неоднозначности авторства за счет работы с полными текстами публикаций, которые доступны в используемой нами коллекции документов поисковой системы Exactus Expert. Данный метод является новым, так как ни в одном из алгоритмов, существующих на данный момент, в качестве признака для кластеризации публикаций не использовалась ключевая лексика, а также ни в одном из подходов не учитывалась семантическая близость публикаций.

Однако данные признаки, очевидно, являются существенными при сравнении работ. В связи с этим ожидается, что включение указанных признаков, повысит точность сопоставления публикаций с соответствующими им авторами.

Благодарности

Публикация подготовлена при поддержке Программы РУДН «5-100» и при финансовой поддержке РФФИ в рамках научных проектов № 12-34-56789 и № 12-34-56789.

Литература

1. Tran H. N., Huynh T. and Do T. Author name disambiguation by using deep neural network // Asian Conference on Intelligent Information and Database Systems. — Springer, 2014. — Pp. 123–132.
2. Han D., Liu S., Hu Y., Wang B. and Sun Y. Elm-based name disambiguation in bibliography // World Wide Web. — 2015. — Vol. 18 (2). — Pp. 253–263.
3. Wu H., Li B., Pei Y. and He J. Unsupervised author disambiguation using Dempster-Shafer theory // Scientometrics. — 2014. — no. 101(3). — Pp. 1955–1972.
4. Liu Y., Tang Y. Network based framework for author name disambiguation applications // International Journal of u- and e-Service. Science and Technology. — 2015. — no. 8(9). — Pp. 75–82.
5. Amancio D.R., Jr. Oliveira O.N. and Costa L.D.F. Topological-collaborative approach for disambiguating authors names in collaborative networks // Scientometrics. — 2015. — Vol. 102 (1). — Pp. 465–485.
6. On B.-W., Elmacioglu E., Lee D., Kang J. and Pei J. Improving grouped-entity resolution using quasi-cliques // ICDM, 2006. — Pp. 1008–1015.
7. Суворов П. Е., Соченков И. В. Определение связанности научно-технических документов на основе характеристики тематической значимости // Искусственный интеллект и принятие решений. — 2013. — №. 1. — С. 33–40.

UDC 004.8

Author's Name Disambiguation on the Document Collection of Exactus Expert

D. S. Novikova

*Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: dsnovikova@inbox.ru

The present work consists of the review of approaches and methods in the field of author's name disambiguation and conclusions about perspective of their usage. Also in the work you can find an information about a direction of our future investigation in the indicated field on the base of the document collection of Exactus Expert.

Key words and phrases: author's name disambiguation.

УДК 004.4

Использование статистических характеристик сигнала в построении аудиоидентификатора музыкальных произведений для поиска нечетких соответствий

А. В. Павлюкова

*Институт системного анализа и управления
Государственный университет «Дубна»
ул. Университетская, д. 19, Дубна, Московская область, Россия, 141980*

Email: anna.pavlyukova@inbox.ru

В работе представлен подход к построению аудиоидентификатора, позволяющий решить задачу поиска нечетких совпадений среди аудио фрагментов различных музыкальных произведений.

Ключевые слова: аудиоидентификатор, статистические характеристики сигнала, поиск нечетких совпадений, информационные технологии.

1. Введение

В последнее время научное сообщество и многие коммерческие компании проявляют большой интерес к задачам распознавания образов. В число таких задач входит и задача построения аудиоидентификатора, осуществляющего поиск исходного музыкального произведения по предоставленному аудиофрагменту.

Наиболее известными зарубежными коммерческими проектами, занимающимися данной задачей, являются Shazam и Spotify. Также существует целый ряд менее известных сервисов. В России данной задачей впервые занялась компания «Яндекс», который впоследствии создал отдельный сервис по поиску музыкальных произведений. За последние несколько лет были разработаны достаточно эффективные алгоритмы поиска четкого соответствия полученного аудиофрагмента и исходного музыкального произведения.

Большинство из них базируются на принципе сравнения «отпечатков» аудио-файлов [1]. Посредством преобразования Фурье исходный сигнал представляют в частотно-временной плоскости: строится спектрограмма [2]. Выбираются пики спектрограмм — частотно-временные точки с наиболее высоким содержанием энергии относительно соседних точек. Такие пики складываются в так называемые «карты созвездий», по которым затем осуществляется поиск [3].

На данный момент существует большое количество модификаций данного алгоритма, включая описанную нами модификацию с применением вейвлет-преобразования. Они весьма эффективны для задач поиска четкого соответствия среди музыкальных произведений. Этот подход позволит получать качественно новую выборку, которая будет в большей степени соответствовать музыкальным вкусам пользователей. Многие из существующих сервисов предоставляют возможность подбора музыкальных рекомендаций, но почти все они базируются на полученных от пользователей статистических данных. Такой подход не способен обеспечить высокое качество получаемых рекомендаций.

2. Основная часть

В этой работе мы хотим предложить подход, который позволил бы реализовать сервис по составлению рекомендаций на основе поиска нечеткого соответствия среди музыкальных произведений. Этот подход позволит получать качественно новую выборку, которая будет в большей степени соответствовать музыкальным вкусам пользователей.

Для этого мы предлагаем обратиться к статистическим характеристикам сигнала. Мы будем использовать статистические характеристики сигнала для анализа

представленной композиции и поиска нечетких соответствий между музыкальными произведениями в собранной нами базе и исходным аудиофрагментом.

Для этого рассчитаем следующие статистические моменты: среднее, стандартное отклонение, асимметрию и эксцесс. Они определяются следующими формулами:

$$\begin{aligned}\mu_k &= E[(X - \mu)^k] = \int_{-\infty}^{\infty} (x - \mu)^k P(x) dx, \\ \sigma^k &= (\sqrt{E[(X - \mu)^2]})^k, \\ \gamma_1 &= E\left[\left(\frac{X - \mu}{\sigma}\right)^3\right] = \frac{\mu_3}{\sigma_3} = \frac{E[(X - \mu)^3]}{(E[(X - \mu)^2])^{3/2}}, \\ Kurt[X] &= E\left[\left(\frac{X - \mu}{\sigma}\right)^4\right] = \frac{\mu_4}{\sigma_4} = \frac{E[(X - \mu)^4]}{(E[(X - \mu)^2])^2},\end{aligned}$$

Рассчитав для каждого сигнала данные характеристики, а также суммарную энергию фурье-спектра и его первой производной, построим из полученных значений вектор представления. Для более глубокого исследования сигнала мы применим кратномасштабный анализ, разложив сигнал по различным масштабам, то есть осуществив его ресемплинг. Так как на разных масштабах статистические характеристики проявляются по-разному, такой подход обеспечит нам высокий уровень детализации.

Для поиска похожих жанровых произведений возьмем евклидову метрику и сравним вектора. В итоге отберем 10 ближайших по метрике произведений, которые и представляют релевантную выборку. Такой подход позволяет нам добиться почти безошибочного определения различных аранжировок исследуемого музыкального произведения. Аранжировки имеют наименьшее значение метрики относительно других композиций, что позволяет нам с легкостью выделять их среди других музыкальных треков. Обычно 4 из 10 подобранных композиций приходятся на аранжировки исходного музыкального произведения.

Помимо поиска аранжировок предоставленного для анализа музыкального произведения, подобный подход позволяет нам также получить релевантную выборку по музыкальным жанрам. При анализе статистических характеристик 5 из 10 подобранных композиций с наименьшей метрикой являются композициями, представленными в одном жанре. Это дает нам право полагать соответствие полученных музыкальных композиций вкусовым предпочтениям пользователя. При тестировании данного подхода на базе в 700 композиций, в 93% случаев алгоритм находил все аранжировки исходного фрагмента, которые находились в базе, и в 74% случаев мы получали подборку схожих композиций в одном жанре.

3. Заключение

Автоматизированная классификация жанра музыкальных композиций и выделение аранжировок исходного музыкального произведения является первым этапом развития алгоритма классификации музыкальных фрагментов. Наша методика, продемонстрировавшая хорошую результативность, может быть развита и использована в дальнейшем для проведения более глубокого анализа, в том числе в качестве базиса для построения полноценного рекомендательного сервиса по подбору музыкальных композиций.

Литература

1. Wang A., An Industrial-Strength Audio Search Algorithm // Proc. 2003 ISMIR International Symposium on Music Information Retrieval, Baltimore, MD, oct. 2003.
2. Ellis D. Robust Landmark-Based Audio Fingerprinting, 2009.
3. Haitsma J., Kalker T. A Highly Robust Audio Fingerprinting System, 2002.

UDC 004.4

Using the Statistical Characteristics of a Signal in the Construction of an Audio Identifier of Musical Works to Search for Fuzzy Matches

A. V. Pavlyukova

*Department of Systems analysis and management
Dubna State University
19 Universitetskaya str., Dubna, Moscow region, 141980, Russian Federation*

Email: anna.pavlyukova@inbox.ru

In this paper, we present an approach to constructing an audio identifier that allows us to solve the problem of finding fuzzy coincidences among the audio fragments of various musical compositions.

Key words and phrases: audio identifier, statistical characteristics of the signal, search for fuzzy coincidences, information technologies.

УДК 004.8

Разработка методов моделирования динамики изменения ишемической болезни мозга путем применения алгоритмов 3d моделирования и морфинга изображений

А. И. Пальчевский*[†], А. И. Молодченков*[‡]

** Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

*† ООО «Технологии системного анализа»
проспект 60-летия октября, д. 9, Москва, Россия, 117312*

*‡ Лаборатория «Интеллектуальные технологии и система»
Федеральный исследовательский центр «Информатика и управление» РАН
проспект 60-летия октября, д. 9, Москва, Россия, 117312*

Email: apalchev@gmail.com, aim@tesya.ru

Ишемический инсульт занимает ведущее место в структуре общемировой смертности и инвалидизации. Основными методами инструментальной диагностики ишемического инсульта головного мозга на сегодняшний момент являются компьютерная томография (КТ) и магнитно-резонансная томография (МРТ). В статье описан метод автоматической обработки МРТ данных для диагностики и прогнозирования развития ишемического инсульта.

Ключевые слова: компьютерные науки, информационные технологии, машинное обучение, искусственный интеллект, медицина, МРТ.

1. Введение

Ишемический инсульт занимает ведущее место в структуре общемировой смертности и инвалидизации. Основными методами лечения данной патологии являются методы реперфузии, которые можно выполнить только в период «терапевтического окна» продолжительность которого не более 8 часов от момента появления симптоматики [1]. В связи с этим быстрая и точная диагностика данной патологии является необходимой для осуществления грамотного лечения пациентов. Основными методами инструментальной диагностики ишемического инсульта головного мозга на сегодняшний момент являются компьютерная томография (КТ) и магнитно-резонансная томография (МРТ) [2].

На сегодняшний день анализ МРТ изображений головного мозга пациентов с подозрением на ишемический инсульт, а также экспериментальных животных, чаще всего производится врачами и учеными вручную. Автоматизация метода обработки МРТ данных [3, 4] является актуальной задачей для современной медицины и биологии. Автоматическая сегментация зоны поражения мозга при ишемическом инсульте позволяет существенно ускорить диагностику и оценку объема поражения, что в свою очередь влияет на тактику лечения пациента. Данный подход также позволяет уменьшить субъективность измерения объема, которая присуща ручной сегментации.

2. Описание метода

Набор данных был предоставлен ФГБОУ ВО РНИМУ им. Н. И. Пирогова Минздрава России. Набор данных включает в себя серию снимков 15 крыс в различных режимах работы томографа.

В ходе реализации алгоритма, обучающий набор данных был обработан для увеличения точности обучения модели нейронной сети. В число данных алгоритмов

для обработки входят: Метод Отсу [5], аффинные преобразования. Далее данные с помощью программного решения с открытым исходным кодом dcm2niix [6] конвертируются в формат Nifti [7].

Реализацию алгоритма можно разделить на несколько стадий: Первый этап — предобработка исходных изображений. Второй этап — разделение набора данных по временному признаку т.е. до поражения, сразу после, двенадцать часов, сутки. Третий этап — выделение мозга и ишемического поражения с помощью нейронной сети. Четвертый этап — построение вероятностной модели развития ишемического поражения на полученных снимках. Пятый этап — построение трехмерной модели и применение методов морфинга для визуализации прогнозирования динамики изменения области поражения. На полученных данных с помощью сверточных нейронных сетей построена модель выделения мозга и ишемического поражения. Точность построенной модели составила 0.9.

Таблица 1

Точность на тестовых данных

Номер	<i>Precision</i>	<i>Recall</i>	<i>f1 – score</i>	Support
CN	0.93	0.93	0.94	159
MCI	0.69	0.79	0.87	157
AD	0.77	0.77	0.86	153
avg / total	0.82	0.85	0.91	469

Для визуализации данных использовался плагин NifTI Analyze [8]. В данном программном решении реализованы алгоритмы построения трехмерных моделей. Для построения модели оценки развития ишемического поражения использовались данные МРТ снимков в динамике. Для визуализации динамики используются алгоритмы морфинга.

3. Заключение

Разработан метод оценки динамики развития ишемического поражения. На данный момент реализован не в полной мере, ведутся работы по программной реализации предсказания динамики развития ишемии головного мозга. Этот метод состоит из следующих этапов: выделение мозга и ишемического поражения на МРТ снимках, преобразование данных для построения 3D модели, оценка динамики изменения на основе данных крыс, визуализация этой динамики.

Благодарности

Работа выполнена в рамках проекта РФФИ № 17-29-07045. «Разработка алгоритмов прогнозирования развития экспериментального ишемического поражения головного мозга на основании интеллектуального анализа данных магнитно-резонансной томографии».

Литература

1. Powers W.J.,Derdeyn C.P., Biller J., Coffey C.S. 2015 American Heart Association / American Stroke Association Focused Update of the 2013 Guidelines for the Early Management of Patients With Acute Ischemic Stroke Regarding Endovascular Treatment: A Guideline for Healthcare Professionals From the American Heart Association/American Stroke Association. [Электронный ресурс]. URL: <https://www.ncbi.nlm.nih.gov/pubmed/26123479> (дата обращения: 12.03.2018).

2. Марусина М. Я., Казначеева А. О. Современные виды томографии. Учебное пособие. — СПб: СПбГУ ИТМО, 2006. — 132 с.
3. Rekić I., Allasonnière S., Carpenter T.K., Wardlaw J.M. Medical image analysis methods in MR/CT-imaged acute-subacute ischemic stroke lesion: Segmentation, prediction and insights into dynamic evolution simulation models. A critical appraisal. *NeuroImage: Clinical*. 2012;1(1):164-178. DOI: 10.1016/j.nic.2012.10.003..
4. Sridharan R., Dalca A.V., Fitzpatrick K.M., et al. Quantification and Analysis of Large Multimodal Clinical Image Studies: Application to Stroke. Multimodal brain image analysis: third International Workshop, MBIA 2013, held in conjunction with MICCAI 2013, Nagoya, Japan, September 22, 2013: proceedings / Li Shen, Tianming Liu, Pew-Thian Yap, Heng Huang, Dinggang Shen, Carl-Fre. 2013;8159:18-30. DOI: 10.1007/978-3-319-02126-3_3..
5. Otsu N., A threshold selection method from gray-level his-tograms // *IEEE Transactions on Systems, Man, and Cybernetics*. — 1979. — Vol. 9, Issue 1. — Pp. 62–66. DOI: 0.1109/TSMC.1979.4310076..
6. rordenlab, Dcm2niix, [Электронный ресурс]. URL: <https://github.com/rordenlab/dcm2niix> (дата обращения: 06.03.2018).
7. Nifti – Neuroimaging Informatics Technology Initiatives, [Электронный ресурс]. URL: <http://nifti.nimh.nih.gov> (дата обращения: 06.03.2018)
8. Williams G., NIFTI Input/Output [Электронный ресурс]. <https://imagej.nih.gov/ij/plugins/nifti.html>.

UDC 004.8

Development of Methods Model of Dynamics of Change of a Chemical Disease of a Brain by Application of the Algorithm of 3D Model and Moving Images

A. I. Palchevskii*[†], A. I. Molodchenkov*[‡]

* *Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

[†] *Technologies for Systems Analysis LLC
Prospekt 60-letia Otechberya, Moscow, 117312, Russian Federation*

[‡] *Laboratory "Intelligent Technologies and Systems"
Federal Research Center "Computer Science and Control" of RAS
Prospekt 60-letia Otechberya, Moscow, 117312, Russian Federation*

Email: apalchev@gmail.com, aim@tesya.ru

Ischemic stroke occupies a leading place in the structure of global mortality and disability. The main methods of instrumental diagnosis of cerebral ischemic stroke at the moment are Computer Tomography (CT) and Magnetic Resonance Tomography (MRI). The article describes the method of automatic processing of MRI data for the diagnosis and prediction of the development of ischemic stroke.

Key words and phrases: computer science, information technologies, conference proceedings, artificial intelligence, machine learning, medicine, MRI.

УДК 519.689.6

Алгоритмы удаления невидимых линий при построении изображений трехмерных тел

К. С. Паршина

*Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: ksushap96@mail.ru

Предметом данного исследования является поиск наиболее эффективного метода построения трехмерного объекта путем удаления невидимых линий в каркасной модели. В работе продемонстрирована классификация различных алгоритмов. В результате данного исследования было сгенерировано итоговое изображение с помощью одного из алгоритмов на языке программирования C++.

Ключевые слова: компьютерная графика, каркасная модель, алгоритм, методы, построение, невидимые линии, трехмерный объект.

1. Введение

В век информационных технологий сфера применения компьютерной графики постоянно расширяется, ведь она позволяет наглядно демонстрировать человеку результат машинных вычислений [3]. Поскольку задача удаления невидимых линий широко применима в компьютерной графике, существует несколько методов её реализации. Выбор хорошего алгоритма и соответствующей структуры данных является наиболее важным для повышения эффективности разрабатываемой программы.

Целью данного исследования является получение навыков работы с трехмерными объектами. Задача заключается в том, чтобы корректно отобразить трехмерное тело в экранном пространстве, удаляя невидимые линии.

2. Основная часть

Главной целью в проектировании алгоритма является выполнение задачи с наименьшим затраченным временем, объемом памяти и пригодность для работы со сложными фигурами [2]. Несмотря на изобилие решений, невозможно выделить единственное и оптимальное. Далее рассмотрим некоторые из них более подробно. Алгоритмы делятся на две категории: общие и специфические. К общим задачам относятся:

- 1) Метод перебора, включающий в себя попарное сравнение ребер и граней. Для каждого ребра происходит проверка на пересечение с каждой плоскостью. Преимущество такого подхода заключается в разрешении задач вне зависимости от количества объектов на сцене и наличия криволинейных плоскостей.
- 2) Алгоритм «художника», заключающийся в построении дальних элементов сцены и последующем закрашивании. Так же он называется алгоритмом сортировки по глубине, поскольку идея этого метода заключается в упорядочивании плоскостей по отдаленности от наблюдателя E и последующем закрашивании от дальних поверхностей к ближним.
- 3) Алгоритм «z-буфера», в котором реализован поиск такого элемента (пикселя), который находится ближе всего к точке наблюдения E. Главной причиной использования данного алгоритма является понятность его идеи, однако он имеет весомый недостаток: при его реализации программе необходимо использовать большой объем памяти. После работы алгоритма получим только те значения z пикселей, которые находятся ближе к наблюдателю.

К специфическим задачам относятся:

- 1) Удаление нелицевых граней у выпуклых тел;
- 2) Удаление невидимых граней в случае, когда часть сцены, содержащая точку наблюдения E , загораживает её другую часть.

Алгоритмы можно сортировать по удалению различных элементов:

- 1) Невидимых линий.
- 2) Невидимых поверхностей.
- 3) Невидимых граней.

Каждый алгоритм служит для решения какой-то конкретной задачи. Для моделирования процессов реального времени главным критерием является скорость исполнения, а для задач построения моделей в компьютерной графике — качество преобразования сложных моделей. Например, алгоритм «z-буфера» эффективен при работе с такими геометрическими фигурами как усеченная пирамида или икосаэдр.

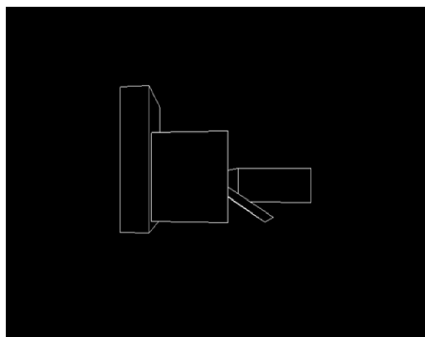


Рис. 1. Пример работы алгоритма. Вид сбоку

Рассмотрим алгоритм, реализованный методом перебора на языке программирования C++ [4]. Разобьём фигуру на треугольные плоскости. Для корректного отображения видимых частей текущего ребра будем смотреть, не закрывает ли его плоскость треугольника. Данный алгоритм заключается в проведении тестов 1–6, описывающих все случаи расположения прямой и плоскости [1]. Для проверки используется пирамида, образованная рассматриваемой плоскостью и точкой, в которой расположена камера:

- 1) Прямая лежит перед плоскостью или на ней.
- 2) Прямая полностью лежит вне пирамиды.
- 3) Прямая и плоскость имеют точку пересечения вне этой плоскости.
- 4) Прямая лежит позади треугольной плоскости или совпадает с её ребром.
- 5) Прямая пересекает плоскость в точке, расположенной перед ней.
- 6) Прямая находится позади плоскости и частично загораживается ею.

3. Заключение

В ходе выполнения поставленной задачи были рассмотрены различные методы удаления невидимых линий и описаны шесть тестов разработанного алгоритма. Качество результата его работы продемонстрировано на примере иллюстрации. Целью дальнейших исследований является модернизация использованного алгоритма и увеличение эффективности его работы.

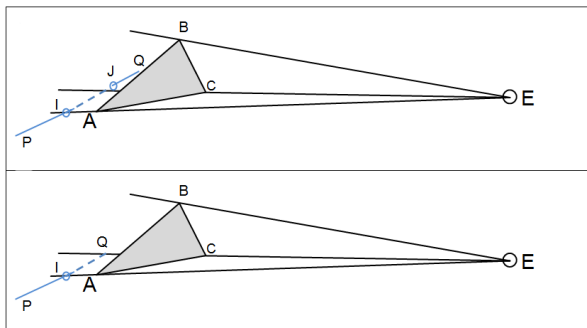


Рис. 2. Расположение прямой и плоскости. Шестой тест

Литература

1. Ammeral L. Programming Principles in Computer Graphics, John Wiley and Sons, 1992.
2. Lamothe A. Tricks of the Windows Game Programming Gurus, Second edition, SAMS, 1999.
3. Dyagterev V. M., Computer Graphics, Publishing Center «Akademiya», 2011.
4. Shikin E. V., Boreskov A. V. Computer graphics. Polygonal models Publishing house «Dialog-MEPhI», 2005.

UDC 519.689.6

Algorithms for Removing Invisible Lines in Constructing Images of Three-Dimensional Objects

K. S. Parshina

*Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: ksushap96@mail.ru

The subject of this study is the search for the most effective method to issue the hidden line removal problem. The classification of various algorithms has been demonstrated to present information about them. As a result of this study, the final image was generated using one of the algorithms in the C++ programming language.

Key words and phrases: computer graphics, wireframe model, algorithm, methods, construction, invisible lines, three-dimensional object.

УДК 004.4

Оценка соответствия экземпляра процесса его обобщенной схеме

О. Д. Перетяtko*, А. И. Молодченков*†

* Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198

† Отдел «Интеллектуальные технологии и системы»
Федеральный исследовательский центр «Информатика и управление» РАН
Пр-т 60-летия Октября, д. 9, Москва, Россия, 117312

Email: peretyatkooleg@gmail.com, aim@tesyan.ru

Для управления процессами организации используют системы моделирования и управления бизнес-процессами. Как показывает практика далеко не всегда реальные процессы соответствуют их обобщенным моделям. Поэтому в 2011 году профессором Виллом ван дер Аалстом был разработан метод оценки соответствия носящий название Token Replay. Впервые данный метод был представлен в книге «Process mining». Но данный метод помогает узнать соответствие лишь в моделях, в которых каждый узел соответствует действию. В более сложных моделях узлы могут соответствовать целым процессам. В статье описана методика оценки соответствия экземпляра процесса его обобщенной схеме.

Ключевые слова: process mining, token replay, BPM, Petri Nets, Workflow Nets, Correspondence metrics.

1. Введение

В современных компаниях для управления процессами, протекающими в них используются системы управления бизнес-процессами. В этих системах специалисты строят модели процессов «как должно быть». Однако далеко не всегда реальные процессы соответствуют их обобщенной схеме. При этом сами модели процессов бывают разного уровня сложности. Бывают модели, в которых каждый узел соответствует действию. В более сложных моделях узлы могут соответствовать целым процессам.

В области анализа (process mining) процессов разрабатываются как методы автоматического синтеза моделей процессов, так и их анализа. Для решения задачи соответствия экземпляра процесса его обобщенной схеме, профессором Wil M.P. van der Aalst был предложен метод Token Replay и выведена соответствующая формула (1). Однако эта формула может быть применена только в случаях, когда узлы модели соответствуют действиям. В статье предложена модификация этой формулы, которая работает и в тех случаях, когда узлы модели соответствуют подпроцессам.

2. Оценка соответствия процесса его обобщенной схеме

2.1. Методика Token Replay

Process mining [1] — это общее название методов и средств для обнаружения анализа и усовершенствования процессов в информационных системах или бизнес-процессов на основании изучения системных данных о выполненных операциях в системе.

Так же для объяснения метода Token Replay требуется ввести определение методики Replay.

Методика Replay [1] — метод проверки соответствия, которая был впервые упомянут профессором Виллом ван дер Аалстом (prof. Wil van der Aalst) в книге “Process mining” выпущенной в 2011 году. Смысл этой методики заключается в

том, что сначала на вход подается схема бизнес-процесса. Далее либо вручную, либо с помощью программных методов журнал событий заполняется различными экземплярами бизнес-процесса.

Пусть есть некоторый бизнес-процесс, на которой была применена методика Replay. Модель процесса описывается в виде Workflow сети — сети Петри с одним началом и концом. Тогда методика Token Replay [2] заключается в том, чтобы, воспроизводить последовательности действий из лога, но при каждом шаге воспроизведения бизнес-процесса метод оставляет некоторые метки или «токены» (из чего собственно и последовало название «Token Replay»).

Метки бывают четырёх типов: метка типа p — произведённые токены, s — потребляемые метки, m — утраченные токены и r — восстановленные метки. Более подробно о том как работают данные метрики расписано в книге “Process Mining” [1] в главе «7.2 Token Replay». Оценка соответствия экземпляра процесса его обобщенной схеме производится по формуле:

$$Fitness(scheme, log) = \frac{1}{2} \left(1 - \frac{m}{c}\right) + \frac{1}{2} \left(1 - \frac{r}{p}\right). \quad (1)$$

Fitness может принимать значения от 0 (экземпляр процесса полностью не соответствует его обобщённой схеме) до 1 (экземпляр процесса полностью соответствует его обобщённой схеме).

2.2. Доработка метода Token Replay

Но такая формула подходит не всегда. Для случаев, когда узел модели соответствует подпроцессу, необходимо производить оценку соответствия внутри этого узла. Поэтому формула (1) была расширена на этот случай и приняла следующий вид:

$$Fitness(scheme, log) = \frac{\frac{1}{2} \left(1 - \frac{m}{c}\right) + \frac{1}{2} \left(1 - \frac{r}{p}\right) + \sum_{i=1}^n Fitness(schema_i, Log_i)}{n + 1}.$$

В данной формуле : n -количество узлов-подпроцессов, $schema_i$ — схема узла подпроцесса, Log_i — часть лог файла с подпроцессом i в узле. Если же вдруг в неком процессе не будет вложенных подпроцессов, то есть $n = 0$, значит все узлы это действия то формула будет аналогична формуле (1).

3. Заключение

В работе была представлена модификация формулы, предложенной Prof.Wil M.P. van der Aalst, для проверки соответствия экземпляра процесса его обобщенной схеме. В отличие от первоначальной формулы, предложенная авторами формула может быть использована в случаях, когда узлы процесса соответствуют как действиям, так и подпроцессам.

Благодарности

Работа выполнена при поддержке гранта РФФИ № 16-29-12839. «Разработка моделей, методов и инструментальных средств для синтеза оптимизированных технологических цепочек и технологических процессов на основе интегрированных баз знаний и интеллектуальных технологий автоматической генерации и оценки планов».

Литература

1. Prof. Wil M.P. van der Aalst. Process Mining. Discovery, Conformance and Enhancement of Business Processes, 2011.
2. Prof. Wil M. P. van der Aalst. Coursera Token Based Replay, 2012.

UDC 004.4

Conformance Checking of a Process Instance to its General Scheme

O. D. Peretyatko*, A. I. Molodchenkov*[†]

** Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

*[†] Department "Intelligent Technologies and Systems"
Federal Research Center "Computer science and control" of RAS
9, Pr-t 60 years of October, Moscow, 11731, Russian Federation*

Email: peretyatkooleg@gmail.com, aim@tesyan.ru

Organizations use systems of modelling and business process management for managing of processes. Experience shows that real processes do not always comply with generalized models. Therefore, in 2011 professor Wil van der Aalst developed the method of assessment of compliance which is called Token Reply. For the first time this method was presented in the book "Process mining". However, this method helps to know the complying only in models where every node matches the act. In more complex models nodes could match the whole processes. The article describes the method of assessment of compliance of the process's exemplar to its generalized schemes.

Key words and phrases: process mining, token replay, BPM, Petri Nets, Workflow Nets, Correspondence metrics.

УДК 519.179.1:510.51

Векторная задача о сочетаниях на гиперграфах

С. И. Салпагаров, Л. М. Гончаров, А. М. Мардашев

*Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклухо-Маклая, д. 6, Москва, Россия, 117198*

Email: salpagarov_si@rudn.university, magicmight2008@gmail.com, mardashev_an@rudn.university

К настоящему времени известен весьма ограниченный перечень задач на гиперграфах и не получила достаточного освещения технология решения таких задач. Под технологией решения задачи понимается четко описанная система действий, выполняемых при ее решении. При этом основной частью этой технологии является алгоритмическая. В данной работе сформулирована математическая постановка векторной задачи о сочетаниях на l -однородных l -дольных гиперграфах. Получена верхняя оценка мощности множества допустимых решений как для задач с равномоными долями, так и для задач с долями разной мощности. Для решения этой задачи построен алгоритм. Первая часть представленного алгоритма нахождения сочетаний l -дольных l -однородных гиперграфах, основана на классическом алгоритме нахождения паросочетаний в двудольном графе. Нахождение допустимых решений задачи выполняется без учета значений весов ребер заданного гиперграфа. Во второй части алгоритма из допустимых решений производится построение множества паретооптимальных решений векторной задачи.

Ключевые слова: гиперграф, многокритериальность, векторная несоравность, паретооптимальность, рекурсивный алгоритм.

1. Введение

В отличие от графовых [1, 2], гиперграфовые постановки векторных (многокритериальных) задач в значительной мере упрощают процесс моделирования реальных объектов и процессов [2, 3]. Так, например, классическая “задача о назначениях”, решение которой представляет собой поиск минимальной (максимальной) суммы весов пересекающихся ребер во взвешенном двудольном графе, может быть обобщена и представлять задачу поиска сочетаний на многодольном гиперграфе.

Преимущество использования теории гиперграфов состоит в том, что ребра гиперграфа позволяют описывать отношение над произвольным множеством объектов, тем самым упрощая моделирование тех задач, в которых невозможно математически задать однозначную зависимость весов на основании бинарных отношений, как в теории графов [1]. При этом следует отметить, что такого типа задачи относятся к классу NP-трудных задач [4], для которых даже в однокритериальной постановке отсутствуют как точные, так и приближенные алгоритмы ее решения.

Для решения сформулированной задачи в работе предложен точный алгоритм нахождения оптимального по Парето множества допустимых решений [5]. Также получена оценка максимальной мощности множества допустимых решений (МДР).

2. Основная часть

Условимся, что недостающие определения по теории гиперграфов можно найти в [2, 3]. Гиперграфом $G = (V, E)$ называется такая пара множеств (V, E) , где V представлено конечным непустым множеством, а E — семейством подмножеств, образованных из множества V . $V = \{v\}$ — множество вершин гиперграфа, а $E = \{e\}$ — множество его ребер. Если пара вершин принадлежит одному ребру, такие вершины называются смежными. Число вершин в ребре называется степенью этого ребра. Сочетание в графе G — это такое подмножество множества ребер E , в котором пересечение каждой пары ребер — пустое множество, то есть, ребра не имеют общих

вершин. Если степень каждого ребра равна l , то такой гиперграф называется l -однородным. Если множество вершин гиперграфа разделено на l подмножеств, в которых каждая пара вершин не является смежной и в любом ребре каждая пара ребер принадлежит различным подмножествам, то гиперграф называется l -дольным гиперграфом. Если каждая пара вершин l -дольного гиперграфа, принадлежащая различным долям смежна, такой гиперграф называют полным l -дольным. Ребро $e \in E$ гиперграфа G называется m -взвешенным, если ему поставлена в соответствие последовательность чисел $w_v(e) \geq 0$, $v = 1, 2, \dots, m$. Гиперграф называется m -взвешенным, если каждое его ребро является m -взвешенным. Качество допустимых решений $x \in X$ оценивается векторной целевой функцией

$$F(x) = (F_1(x), F_2(x), \dots, F_m(x)), \quad (1)$$

критерии которой имеют вид *MAXSUM* [6]:

$$F_v(x) = \sum_{e \in x} w_v(e) \rightarrow \max, \quad v = 1, 2, \dots, m.$$

3. Описание работы алгоритма

Исходными данными алгоритма являются: l -дольный l -однородный m -взвешенный гиперграф $G = (V, E) = (V_1, \dots, V_l, E)$, $|V_{\min}|$ — мощность меньшей доли гиперграфа, $|V_{\min}| = \min\{|V_1|, |V_2|, \dots, |V_l|\}$. Результат работы рассматриваемого алгоритма будет представлен парой множеств (X^0, W) , поставленных в однозначное соответствие. X^0 — множество оптимальных по Парето решений и W — множество весов в виде (w_1, \dots, w_m) . Допустимым решением называется такое сочетание, которое содержит все вершины доли V_{\min} .

Алгоритм разделен на две части, решающие отдельные подзадачи. Первая часть алгоритма связана с нахождением множества допустимых решений МДР без учета весов ребер. Вторая часть алгоритма на основании весов выделяет из МДР подмножество оптимальных по Парето решений. В алгоритме используются промежуточные множества T, R . Построение каждого допустимого решения основано на рекурсивном выделении из заданного множества ребер E подмножества R , которое дополняется ребром из $T \subseteq E$ на каждом уровне рекурсии таким, что оно несмежно с любым ребром из R . Если на рассматриваемом шаге рекурсии R не достигает мощности $|V_{\min}|$, рекурсивный поиск продолжается и в качестве T на следующий шаг передается T' — подмножество ребер из T , таких что $T' \subseteq \{T \setminus \{t_j\}\} : \forall (t_j, t'_k), t'_k \in T', t'_k \cap t_j = \{\emptyset\}$. Если на рассматриваемом шаге $T = \{\emptyset\}$, построение текущего допустимого решения завершается.

В работе была получена оценка максимальной мощности МДР [7]:

$$|X(G)| \leq \prod_{j=0}^p \prod_{k=1}^l (|V_k| - j), \quad p = |V_{\min}| - 1.$$

Для равнодольного гиперграфа формулу можно представить в виде:

$$|X(G)| \leq \left(\binom{|V|}{l} \right)^{l-1}.$$

Данная граница является достижимой только в случае, когда заданный l -дольный l -однородный гиперграф полный.

Нахождение оптимальных по Парето решений среди допустимых решений строится по следующему принципу. Предполагается, что критерии вида (1) представлены

числами и нормализованы. По вектору весов (w_1, \dots, w_m) каждая пара решений может находиться в четырех состояниях относительно друг друга: первое качественно лучше второго, второе качественно лучше первого, решения качественно эквивалентны и решения векторно несравнимы. Чтобы однозначно определить отношение между парой решений, достаточно знать количество критериев первого решения, качественно превосходящих второе, и количество эквивалентных критериев.

4. Заключение

Сформулирована векторная задача о сочетаниях на однородных многодольных гиперграфах. Получена оценка максимальной мощности МДР, а так же построен алгоритм решения сформулированной задачи. Представленная задача, построение ее математической модели на базе теории гиперграфов, алгоритмы ее решения помогут преодолеть трудности в исследовании дискретных задач управления.

Литература

1. Татт У. Теория графов. — М.: Мир, 1988. — 320 с.
2. Лекции по теории графов / В. А. Емеличев, О. И. Мельников, В. И. Сарванов, Р. И. Тышкевич. — М.: Наука, 1990. — 384 с.
3. Зыков А. А. Гиперграфы // Успехи математических наук. — 1974. — Т. 29, вып. 6. — С. 89–154.
4. Гэри М., Джонсон Д. Вычислительные машины и труднорешаемые задачи. — М.: Мир, 1982. — 416 с.
5. Подиновский В. В., Ногин В. Д. Парето-оптимальные решения многокритериальных задач. — М.: Наука. Главная редакция физико-математической литературы, 1982. — 256 с.
6. Емеличев В. А., Перепелица В. А. Сложность дискретных многокритериальных задач // Дискретная математика. — 1994. — Т. 6, № 1. — С. 3–33.
7. Новиков Ф. А. Дискретная математика для программистов. — СПб.: Питер, 2009. — 384 с.

UDC 519.179.1:510.51

Vectorial Problem on Combinations on Hypergraphs

S. I. Salpagarov, L. M. Goncharov, A. M. Mardashev

*Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: salpagarov_si@rudn.university, magicmight2008@gmail.com, mardashev_am@rudn.university

To date, a very limited list of problems on hypergraphs is known and the technology for solving such problems has not been adequately covered. The technology of solving a problem is understood as a clearly described system of actions performed in its solution. The main part of this technology is algorithmic. In this paper, a mathematical formulation of the vectorial problem of combinations on l -homogeneous l -partite hypergraphs is formulated. An upper bound is obtained for the cardinality of the set of admissible solutions both for problems with equipotential parts and for problems with parts of different cardinality. An algorithm is constructed to solve this problem. The first part of the presented algorithm for finding combinations on l -partite l -homogeneous hypergraphs is based on the classical algorithm for finding matchings in a bipartite graph. The determination of admissible solutions of the problem is performed without taking into account the values of the weights of the edges of a given hypergraph. In the second part of the algorithm, from the admissible solutions, a set of paretooptimal solutions of the vectorial problem is constructed.

Key words and phrases: hypergraph, multicriteria, vector incomparability, Pareto optimality, recursive algorithm.

UDC 519.179.1:654

Multicriteria Model of Distribution of P2P-TV Data Streams on Hypergraphs

S. I. Salpagarov, Yu. D. Isaev

*Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: salpagarov_si@pfur.ru, ydisaev@mail.ru

This article describes the problem of the distribution data streams process in P2P video streaming systems. It also discuss the case of watching TV channels by a lot of users in P2P system. The mathematical model is described in the language of the hypergraph theory in a multi-criterion formulation. The criteria are MINMAX and MINMIN, which allow taking into account the performance indicators such as the time of switching between channels, the interruption of the image and the minimum traffic load at the same time.

Key words and phrases: Hypergraph; combinations on hypergraphs; P2P-TV; distribution of data flows; vector objective function; multicriteria.

1. Introduction

Nowadays peer-to-peer (P2P) file sharing systems are widely used in many areas of human activity. First of all, they are used for data exchange and transmission of streaming content. In these networks, many television channels are viewed simultaneously by a large number of users. A number of studies [1–4] are devoted to the analysis of performance indicators of peer-to-peer video streaming networks (P2PTV). Various methods of research are used, such as the construction and analysis of analytical models [5], which use the apparatus of the theory of exponential queuing networks [6].

The VUD mechanism clearly separates what the user is downloading and viewing, thereby the stability of the multi-channel system and the ability to share resources between channels achieve. Each user is assigned to one or more channels, regardless of what the user is viewing. The user downloads and sends to other users all the data of the channels assigned to him.

The purpose of this study is to build a mathematical model for efficient organization of the distribution of data flows between users of P2P networks for the view-upload decoupling (VUD) scheme. Using this model, it will be possible to find solutions to the problem in which both the switching time between channels and the interruption time of the playback of the channels being watched will be minimal.

2. Main section

The objects for constructing the mathematical model of data transmission are:

1. $U = \{u\}$ — the set of users receiving the service;
2. $C = \{c\}$ — the set of channels available to the user;
3. $F = \{f\}$ — the number of threads to which channels can be allocated.

Each user $u_i \in U$ receives streams $f \in F$ of the channel $c \in C$ from users $u_j \in U$, and $u_i \cap u_j = \emptyset$. The result of this distribution of channel flows between users should be an increase in the level of performance of the broadcasting system, in particular, the quality of service of unpopular channels should not be much worse than the quality of service of channels that are more popular.

The mathematical model considered in this article is based on a 3-fractional 3-homogeneous hypergraph $G = (V, E) = (V_1, V_2, V_3, E)$ [7, 8], which is constructed as follows. Vertices of the first fraction $v \in V_1$, uniquely correspond to the elements of a set of users U . Each vertex $v \in V_1$ corresponding to the user $u \in U$, is assigned the

number $m(v)$, that determined by the number of threads that the user will transmit. Each vertex of the second fraction $v \in V_2$ uniquely corresponds to some element from the set of threads F , distributed to users. Vertexes of the third fraction $v \in V_3$ uniquely correspond to the elements of the set of channels C . To construct a set of edges $E = e$ all possible triples of vertexes consider (v_1, v_2, v_3) so that $v_1 \in V_1, v_2 \in V_2, v_3 \in V_3$. Any such triple is called acceptable, if the user v_1 has an opportunity to stream data thread v_2 of the channel v_3 . The set of all edges $E = e$ is defined as the set of all acceptable triples $e = (v_1, v_2, v_3), v_i \in V_i, i = (\overline{1, 3})$.

In the problem under consideration for a hypergraph $G = (V_1, V_2, V_3, E)$ the following conditions are fulfilled:

1. In each edge $e = (v_1, v_2, v_3) \in E$ a pair of vertexes v_1, v_3 , which are called terminal vertexes for the edge, is allocated;
2. Vertexes $v \in V_2$ are internal, and the set V_2 is consisted of nonempty pairwise disjoint sets $V_2(v_3), v_3 \in V_3$, and each element $v \in V_2(v_3)$ uniquely corresponds to some thread $f \in F$;
3. Terminal vertexes $v_3 \in V_3^Z$ are hanging vertexes (power 1);
4. For each vertex v from V_1 the number $m(v)$ is indicated. The number $m(v)$ is a parameter of the following condition: a star with center at the vertex $v \in V_1$ has a power $r(v) = m(v)$ and $\sum_{v \in V_1} m(v) = |V_3|$.

Let us set $X = X(G) = x$ as the acceptable solutions set of the problem of covering the hypergraph G by stars. Each edge $e \in E$ of the hypergraph $G = (V, E)$ has three weights $w_1(e), w_2(e)$ and $w_3(e)$, where $w_1(e) = f_1(v_1, v_2, v_3)$ is access delay when a user $v_1 \in V_1$ is switching on the channel $v_3 \in V_3$, for the transmission of which the stream $v_2 \in V_2$ was used, $w_2(e) = f_2(v_1, v_2, v_3)$ the interrupt time when the image is played in the same case, and $w_3(e) = f_3(v_1, v_2, v_3)$ is the throughput of the edge $e \in E$.

The quality of the acceptable solutions of this problem $x \in X$ is estimated using the vector objective function $F(x) = (F_1(x), F_2(x), F_3(x))$ with type MINMAX $F_i(x) = \max w_i(e) \rightarrow \min, i = 1, 2$ and type MINMIN $F_i(x) = \min w_i(e) \rightarrow \min, i = 3$, where $F_1(x)$ is the expected level of channel switching time in a given solution x ; $F_2(x)$ is change the interruption of playback in the solution x , and $F_3(x)$ is the minimum traffic load in a given solution x .

3. Conclusions

As a result of the work, a multi-criteria hypergraph model was constructed, which solved the problem of efficient distribution of data threads between users of P2P networks for the VUD scheme. This model allows us to find solutions to the problem for which the switching time between channels and the interruption time of the playback of the watched channels are minimal.

In this article the solution of the problem of optimizing the process of distribution of flows and minimization of traffic load in P2PTV using concepts and objects of the theory of hypergraphs is described. The application of this theory in combination with the elements of the theory of multicriteria optimization makes it possible to take into account in the systemic unity a complex organization of internal interrelations of the elements of the problems under consideration.

Acknowledgments

The authors are grateful to Professor Samuylov K. Ye., and Gaydamaka Yu. V. for their support in writing and publishing this work.

References

1. D. Wu, Y. Liu, K. Ross, Queuing Network Models for Multi-Channel P2P Live Streaming Systems, IEEE INFOCOM. 2009. Pp. 73–81.
2. Yu. V. Gaidamaka, A. K. Samuylov, Analysis of playback continuity for video streaming in Peer to Peer networks with data transfer delays, T-Comm, 2013, no 11, Pp. 77–81, in Russian.
3. A. K. Samuylov, E. V. Bobrikova, A simple fluid model of P2P file sharing network, T-Comm, 2012, no. 7, Pp. 180–184, in Russian.
4. A. Adamu, Yu. V. Gaidamaka, Approximation of a Universal Streaming Probability with the Normal Distribution, Bulletin of Peoples' Friendship University of Russia, Series: Mathematics. Information Sciences. Physics, 2011, no. 3, Pp. 63–68, in Russian.
5. Yu. V. Gaidamaka, E. G. Medvedeva, S. I. Salpagarov, E. V. Bobrikova, Analysis of Model for Multichannel Peer-to-Peer TV Network with View-Upload Decoupling Scheme, RUDN Journal of Mathematics, Informational Sciences and Physics, 2017, no. 2, Pp. 123–132, in Russian.
6. G. P. Basharin, Lekcii po matematicheskoj teorii teletrafika: ucheb. posobie. izd. 3-e, ispr. i dop., M.: RUDN, 2009, 342 p., in Russian.
7. Emelichev V.A., Mel'nikov O.I., Sarvanov V.I., Tyshkevich R.I. Lekcii po teorii grafov, M.: Nauka, 1990, 384 p., in Russian.
8. A. A. Zykov, Gipergrafy, Uspehi Matem. nauk, 1974, Vol. 29, Issue 6, Pp. 89–154, in Russian.
9. S. I. Salpagarov, G. G. Omel'chenko, Modelirovanie na gipergrafah, M.: RUDN, 2010, 64 p, in Russian.

УДК 519.179.1:654

Оптимизационная модель распределения потоков данных P2P-телевидения на гиперграфах

С. И. Салпагаров, Ю. Д. Исаев

*Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклуло-Маклая, д. 6, Москва, Россия, 117198*

Email: salpagarov_si@pfur.ru, ydisaev@mail.ru

В данной статье описана проблема процесса распределения потоков данных в одноранговых сетях передачи потокового видео. Рассмотрен случай просмотра телевизионных каналов множеством пользователей в P2P сетях. Математическая модель описана на языке теории гиперграфов в многокритериальной постановке. Критерии имеют вид MINMAX и MINMIN, которые позволяют одновременно учитывать показатели эффективности по времени переключения между каналами, прерывания изображения и минимальной загрузке трафика.

Ключевые слова: гиперграф; сочетания на гиперграфах; P2P-телевидение; распределение потоков данных; векторная целевая функция; многокритериальность.

УДК 004.8

Разработка метода оценки близости пептидов

Л. Сингх*, А. И. Молодченков^{†‡}

* *Факультет компьютерных наук
Высшая школа экономики
Кочновский проезд, д. 3, Москва, Россия, 125319*

† *Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

‡ *Лаборатория «Интеллектуальные технологии и системы»
Федеральный исследовательский центр «Информатика и Управление» РАН
проспект 60-летия октября, д. 9, Москва, Россия, 117312*

Email: preet.94.kaur@mail.ru, aim@tesyan.ru

Известно, что белок человека имеет уникальную аминокислотную последовательность в пределах генома, по которой его можно идентифицировать. Поэтому, имея количественный протеомный анализ некоторого участка в известном белке, можно определить, содержит ли искомый белок такую же уникальную последовательность. Такая задача до сих пор решается биоинформатиками вручную. В связи с этим, целью данной работы было поставлено автоматизирование процесса анализа биологического образца, который имеет некоторое количество различных белков. В качестве алгоритма предлагается построение модели классификации, на основе некоторого образца с результатами эксперта. Данная работа направлена на анализ протеомных или других био-медицинских данных в сферах биоинформатики и информационных технологий в медицине.

Ключевые слова: компьютерные науки, информационные технологии, искусственный интеллект, машинное обучение, биоинформатика, протеомика, масс-спектрометрия.

1. Введение

На сегодняшний день уже существуют некоторые методы идентификация белка. Например, секвенирование аминокислот белка или более распространенный метод — peptide-mass finger printing [1]. Однако, оба способа неприемлемы при анализе большого числа белков. Более того, они требуют выделение отдельного белка и не позволяют проанализировать образец, содержащий несколько различных белков.

Для поиска определенных белков в образце используется метод количественной протеомики. Результаты представляются в виде хроматограмм, описывающих уникальные пептиды белков из исследуемого образца и пептиды, синтезированные в лабораторных условиях, то есть некоторые стандарты. В настоящий момент эксперт анализирует их визуально путем сравнения идентифицируемого пептида с его стандартом. Это приводит к случайным ошибкам, допущенным экспертом.

Цель данной работы — реализация метода автоматической оценки близости пептидов для минимизации количества экспертных ошибок. Этот метод заключается в сведении данной задачи к созданию модели классификации пар пептидов, стандартного и идентифицируемого, которые делятся на два класса: схожи и не схожи.

2. Основная часть

Одним из методов количественной протеомики является метод мониторинга выбранных реакций (МВР) тандемной масс-спектрометрии во времени [2]. Он осуществляется путем поиска определенного пептида по его массе и его фрагментным ионам. Результатом масс-спектрометрического анализа является изменение интенсивности каждого искомого иона во времени.

Так как белок содержит уникальную аминокислотную последовательность в пределах генома человека, его можно идентифицировать по наличию такого пептида в образце. Чтобы иметь возможность интерпретировать результаты МС анализа, к биологическому материалу добавляются изотопно меченые стандарты, то есть пептиды, синтезированные в лабораторных условиях. Таким образом, с помощью попарного сравнения результатов между известным и искомым белками можно идентифицировать пептид и сам белок.

Чтобы свести количественный анализ пептидов в образце к модели классификации, необходимо собрать выборку. Каждый пептид характеризуется тремя хроматограммами, или изменением интенсивности трех фрагментных ионов пептида во времени. Тогда, объектом выборки будем считать пару пептидов, идентифицируемого и стандартного. А описанием объектов - признаки сходства между тремя хроматограммами.

Примеры хроматограмм приведены на рис. 1.

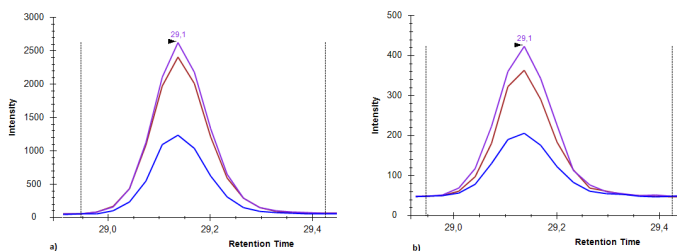


Рис. 1. Хроматограммы синтезированного (а) и идентифицируемого (б) пептидов

Для определения признаков рассмотрим первую и вторую производные каждой хроматограммы, то есть рассчитаем скорость и ускорение изменения интенсивности ионов во времени:

$$f_1 = \frac{dx(t)}{dt}; f_2 = \frac{df_1}{dt} = \frac{d^2x(t)}{dt^2}.$$

Тогда, признаком сходства можно считать корреляцию между скоростями и ускорениями для каждого иона.

Таким образом, была получена выборка из 86 пар пептидов, характеризующихся 6 числовыми признаками. Из них 66% схожих пептидов и 34% — нет.

При построении модели идентификации пептидов были использованы следующие алгоритмы классификации: метод ближайших соседей, дерева решений, метод опорных векторов и бэггинг, использующий комбинацию всех трех предыдущих методов. В табл. 1 представлены результаты работы алгоритмов:

Как показали эксперименты наилучшее качество по F1-мере показал метод ближайших соседей с параметрами: 21 ближайших соседей, метрика Чебышева. Перед обучением модели данные были приведены к стандартному нормальному распределению, с нулевым средним значением и единичным стандартным отклонением. Затем был проведен подбор наилучших параметров по сетке, где 90% обучающей выборки и 10% — тестовой.

3. Заключение

Разработан метод автоматической идентификации белков. Преимущество разработанного метода заключается в том, что он позволяет определить наличие

Таблица 1

Результаты экспериментов

Алгоритм	Значение F1-меры
Метод ближайших соседей	0.84
Деревья решений	0.75
Метод опорных векторов	0.79
Беггинг	0.82

уникальных пептидов в исследуемом образце и таким образом идентифицировать наличие определенных белков в этом образце, не прибегая к более сложным методам, таким как секвенирование и peptide-mass finger printing.

Литература

1. Pappin D. J., Hojrup P., Bleasby A. J., Rapid identification of proteins by peptide-mass fingerprinting, *Curr Biol.*, UK, 1993.
2. Количественная протеомика белков хромосомы 13 в плазме крови человека / А. Т. Копылов, Е. В. Ильгинсонис, О. В. Тихонова, Т. Е. Фарафонова, С. Е. Новикова, В. Г. Згода, Е. А. Пономаренко, академик РАН А. В. Лисица, А. А. Маркин, Б. В. Моруков, А. И. Григорьев, А. И. Арчаков // Доклады академии наук, Федеральное Государственное Унитарное Предприятие «Академический научно-издательский, производственно-полиграфический и книгораспространительский центр «Наука», 2017.

UDC 004.8

Development of Method for Estimation of Peptide Similarity

L. Singh*, A. I. Molodchenkov^{†‡}* Faculty of Computer Science
High School of Economics

3, Kochnovsky Proezd, Moscow, 125319, Russian Federation

† Department of Information Technologies

Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation‡ Laboratory "Intellectual Technologies and Systems"
Federal Research Center "Computer Science and Control" of RAS
Prospekt 60-letia Otechberya, Moscow, 117312, Russian Federation

Email: preet.94.kaur@mail.ru, aim@tesyan.ru

It is known that a human protein has a unique amino acid sequence within the genome, according to which it can be identified. Therefore, having a quantitative proteomic analysis of the region in known protein, it can be determined whether the desired protein contains the same unique sequence. This problem is still solved by bioinformatics manually. In this case, the purpose of this work is to automate the process of analyzing a biological sample that has a number of different proteins. The algorithm is proposed to construct a classification model, based on a sample with the results of an expert. This work is aimed at the analysis of proteomic or other bio-medical data in the fields of bioinformatics and information technology in medicine.

Key words and phrases: computer science, information technologies, artificial intelligence, machine learning, bioinformatics, proteomics, mass spectrometry.

УДК 004.021

Исследование метода классификации новостей на основе дистрибутивной семантики

К. В. Скрынник

*Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: skrinnikkv@gmail.com

В статье рассматривается метод классификации новостей с использованием дистрибутивной семантики текста. Показано влияние признаков, полученных с использованием моделей дистрибутивной семантики, на метрики качества классификации. В качестве набора данных взят «The 20 Newsgroups dataset». Представлены результаты сравнительного анализа популярных классификаторов, таких как Random Forest, KNeighbours, Support Vector Machine, Logistic Regression.

Ключевые слова: классификация текстов, дистрибутивная семантика, SVM, kNN, RF, LR.

1. Введение

В статье исследована задача классификации новостей с использованием методов дистрибутивной семантики. Алгоритм в общих чертах выглядит так. Сначала текстовые данные очищаются от знаков препинания, чисел, стоп-слов и другого "мусора". Затем при помощи инструмента Word2Vec для слов находятся их векторные представления. В дальнейшем текстовые данные, представленные в виде списка векторных представлений входящих в них слов подаются на вход различным классификаторам.

2. Основная часть

2.1. Предварительная подготовка входных данных

На этом этапе задача сводится к получению максимально эффективного векторного представления входных данных. Для этого текстовые данные необходимо нормализовать.

2.1.1. Очистка и нормализация текстов

Все тексты должны содержать лишь символы, образующие слова и разделяющие их пробелы. Поэтому для начала из текста необходимо удалить все числа и знаки препинания. Для выполнения этого рационально использовать регулярные выражения. Далее, с помощью инструмента Cisso, текст очищается от прочего «мусора», такого как адреса электронной почты или ссылки на веб-сайты. После этого слова приводятся к их нормальной форме при помощи инструмента Snowball Stemmer из библиотеки nltk.stem.

Последним шагом на данном этапе необходимо проверить полученный текст на наличие в нем стоп-слов¹.

¹ стоп-слова — часто встречающиеся в текстах слова, не несущие смысловой нагрузки.

2.1.2. Нахождение векторного представления текстов

Теперь входные данные прошли предварительную обработку и для них можно найти векторное представление.

Предварительно необходимо, при помощи инструмента `gensim` загрузить обученную на большом массиве текстовых данных модель `Word2Vec`.

Теперь каждый входной текст разбивается на отдельные слова, для каждого из них, при помощи модели `W2V` находится векторное представление. Полученный список векторов и будет являться векторным представлением исходных данных.

На этом предварительную подготовку входных данных можно считать завершённой.

2.2. Использование классификаторов и анализ результатов

Теперь входные данные можно использовать для обучения и тестирования различных классификаторов. Для этого необходимо разбить их на обучающую и тестовую выборки. Для тестирования в данном случае было выбрано «классическое» разбиение — 80% — для обучения классификаторов — 20% — для тестов.

Для тестирования были выбраны следующие классификаторы - `Random Forest`, `KNeighbours`, `Support Vector Machine`, `Logistic Regression`. Тестирование проводилось для входных данных подготовленных следующими способами:

- с использованием только регулярных выражений;
- с использованием регулярных выражений и общей нормализацией текста;
- с использованием регулярных выражений, общей нормализации текста, а также нормализацией отдельных слов.

Настройки классификаторов:

- `Random Forest`: `n_estimators=300`, `max_depth=100`, `random_state=4`
- `KNeighbours`: `n_neighbors=4`, `random_state=4`
- `SVM`: `C=70`, `tol=0.0001`
- `Logistic Regression`: `C=50`, `random_state=4`

Результаты работы классификаторов для всех тестов можно видеть в табл. 1–3.

Таблица 1

Регулярные выражения

Classifier	Precision	Recall	F-measure	Operating time
Random Forest	0.73	0.72	0.71	64 sec.
KNeighbours	0.75	0.74	0.74	25.2 sec.
SVM	0.76	0.75	0.75	48.9 sec.
LR	0.79	0.79	0.79	39 sec.
Best result:	0.79	0.79	0.79	25.2

Проанализировав полученные данные можно сделать вывод, что наилучший результат показывает логистическая регрессия, обученная на данных, очищенных регулярными выражениями.

3. Заключение

В этой работе была решена задача классификации текстовых данных с использованием методов дистрибутивной семантики. Был выявлен наиболее подходящий для этой задачи классификатор.

Таблица 2

Регулярные выражения, нормализация текста

Classifier	Precision	Recall	F-measure	Operating time
Random Forest	0.72	0.72	0.71	69 sec.
KNeighbours	0.75	0.73	0.73	15.9 sec.
SVM	0.74	0.74	0.74	39.9 sec.
LR	0.77	0.77	0.77	35.3 sec.
Best result:	0.77	0.77	0.77	15.9

Таблица 3

Регулярные выражения, нормализация текста, нормализация слов

Classifier	Precision	Recall	F-measure	Operating time
Random Forest	0.73	0.72	0.72	72 sec.
KNeighbours	0.75	0.73	0.73	12.6 sec.
SVM	0.74	0.74	0.74	40.5 sec.
LR	0.77	0.77	0.77	35.4 sec.
Best result:	0.77	0.77	0.77	12.6

Литература

1. Albishre K., Albathan M., Li Y., Effective 20 newsgroups dataset cleaning // Web Intelligence and Intelligent Agent Technology (WI-IAT). — 2015.
2. Shaikh J., Machine Learning, NLP: Text Classification using scikit-learn, python and NLTK. — 2017.

UDC 004.021

Study of the Method of Classification of News Based on Distributive Semantics

K. V. Skrinnik

*Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: skrinnikkv@gmail.com

This paper is about a method for classifying news using the distributive semantics of the text. The influence of signs, features using models of distributive semantics, on the quality metrics of classification is shown. As a set of data was taken "The 20 Newsgroups dataset". Presented results of a comparative analysis of popular classifiers, such as a Random Forest, KNeighbours, Support Vector Machine, Logistic Regression.

Key words and phrases: text classification, distributive semantic, SVM, kNN, RF, LR.

УДК 004.4

Вычислительно эффективный алгоритм восстановления изображений с использованием полной вариации

К. Тханг Фам*, А. Копылов†, Т. Т. Тхао Чан†

* *Национальный исследовательский университет «Высшая школа экономики»
Москва, Россия*

† *Тульский государственный университет, Тула, Россия*

Email: pacotha@gmail.com

В данной работе представлен эффективный алгоритм устранения шума на изображении с использованием полной вариации. Предлагаемая новая численная схема позволяет одновременно удалять как пуассоновский, так и гауссовский шум, с меньшими вычислительными затратами. Экспериментальные результаты показывают эффективность предлагаемого алгоритма.

Ключевые слова: шумоподавление изображения, гауссовский шум, шум Пуассона, полная вариация, градиентный поток.

1. Введение

В процессе получения и передачи по каналам связи изображение как правило подвергается воздействию шума различной природы. В связи с этим, устранение шума играет важную роль для любых систем анализа цифровых изображений. Важным требованием для процедуры устранения шума является сохранение локальных структур изображения для точного и эффективного последующего анализа. Несмотря на большое количество исследований, посвященных устранению шумов различного типа (гауссовские, пуассоновские, райсовские и т. д.), чаще всего рассматривается шум лишь одного возможного вида. Задача шумоподавления обычно ставится как задача восстановления истинного изображения u поврежденного некоторым шумом η , по наблюдаемому изображению f .

Одним из самых распространенных способов восстановления изображения u является использование метода полной вариации на основе известной ROF модели [1], доказавшей свою эффективность в задаче удаления гауссовского шума. В работе [2], Le и другие предложили на основе ROF модели новую модель, которая способна удалять пуассоновский шум. В этом случае необходимо решить следующую оптимизационную задачу:

$$u^* = \arg \min_u E(u) = \int_{\Omega} |\nabla u| + \beta \int_{\Omega} (u - f \log u), \quad (1)$$

где $\Omega \subset \mathbb{R}^2$, u — функция ограниченной вариации на Ω , $u \in BV(\Omega)$ [2].

Для численного решения (1), авторы в [2] использовали следующую итерационную схему:

$$\frac{u^{(n+1)} - u^{(n)}}{\tau} = \operatorname{div} \left(\frac{\nabla u^{(n)}}{|\nabla u^{(n)}|} \right) - \beta \left(1 - \frac{f}{u^{(n)}} \right),$$

где $n = 1, 2, \dots$, $\nabla u_{j,k} = (u_{j,k}^{(x)}, u_{j,k}^{(y)})$, $u_{j,k}^{(x)} = u_{j+1,k} - u_{j-1,k}$, $u_{j,k}^{(y)} = u_{j,k+1} - u_{j,k-1}$, $|\nabla u| = \sqrt{(u^{(x)})^2 + (u^{(y)})^2 + \varepsilon^2}$, $u_{j,k}^{(xx)} = u_{j+1,k} - 2u_{j,k} + u_{j-1,k}$, $u_{j,k}^{(yy)} = u_{j,k+1} - 2u_{j,k} + u_{j,k-1}$, $u_{j,k}^{(xy)} = u_{j+1,k+1} + u_{j-1,k-1} - u_{j+1,k-1} - u_{j-1,k+1}$, $\operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right) = \frac{u^{(xx)}(u^{(x)})^2 - 2u^{(x)}u^{(y)}u^{(xy)} + u^{(yy)}(u^{(y)})^2}{(u^{(x)})^2 + (u^{(y)})^2 + \varepsilon^2}^{3/2}$, $j = 1, \dots, N$, $k = 1, \dots, M$.

Недавно W. Wang и С. Не [3] предложили другую, более эффективную схему для решения (1):

$$\frac{u^{(n+1)} - u^{(n)}}{\tau} = \operatorname{div} \left(\frac{\nabla u^{(n)}}{|\nabla u^{(n)}|} \right) - \beta \left(1 - \frac{f}{u^{(n+1)}} \right). \quad (2)$$

Рассмотренные модели [1–3] эффективны для восстановления изображения, в случае если присутствует либо гауссовский, либо пуассоновский шум. В работе [4] авторы предложили линейную комбинацию моделей [1, 2] для восстановления изображения со смесью пуассоновского и гауссовского шума:

$$u^* = \arg \min_u E(u) = \int_{\Omega} |\nabla u| + \frac{\lambda}{2} \int_{\Omega} (u - f)^2 + \beta \int_{\Omega} (u - f \log u). \quad (3)$$

Однако, модель (3) не допускает одновременного изменения дисперсии пуассоновского и гауссовского шума. В данной работе авторы использовали модель [2] для создания комбинированной модели. Как показано в [3], найденное решение u^* задачи (3) может включать как положительные, так и отрицательные элементы. В то же время целевая функция (3) содержит $\log(u)$ и такое решение не может быть оптимальным. В этой статье мы сосредоточимся на разработке нового алгоритма, который может восстанавливать изображение со смесью пуассоновского и гауссовского шума на основе новой, более эффективной схемы решения задачи (3), позволяющей избежать проблемы изменения знака решения u . Эксперименты показывают что, предлагаемая схема превосходит сравнительные методы по точности для удаления смешанного гауссовско-пуассоновского шума.

2. Предлагаемая схема для удаления смешанного гауссовско-пуассоновского шума

Метод градиентного спуска для задачи (3) основан на следующем выражении:

$$\frac{\partial u}{\partial t} = \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right) - \beta \left(1 - \frac{f}{u} \right) - \frac{\lambda}{\sigma^2} (u - f),$$

где σ — оценка дисперсии шума методом Иммеркера [5], $\lambda > 0$, $\beta > 0$, $\lambda + \beta = 1$.

Мы предлагаем следующую новую схему для нахождения оптимального решения задачи (3) на основе (2):

$$\frac{u^{(n+1)} - u^{(n)}}{\tau} = \operatorname{div} \left(\frac{\nabla u^{(n)}}{|\nabla u^{(n)}|} \right) - \beta \left(1 - \frac{f}{u^{(n+1)}} \right) - \frac{\lambda}{\sigma^2} (u^{(n+1)} - f), \quad (4)$$

где $\tau \in (0, 1)$, $\lambda > 0$, $\beta > 0$, $\lambda + \beta = 1$.

Уравнение (4) может быть переписано как:

$$a_n \left(u^{(n+1)} \right)^2 + b_n u^{(n+1)} + c_n = 0, \quad (5)$$

где $a_n = \left(1 + \tau \frac{\lambda}{\sigma^2} \right)$, $b_n = - \left(u^{(n)} + \tau \operatorname{div} \left(\frac{\nabla u^{(n)}}{|\nabla u^{(n)}|} \right) - \tau \beta + \tau f \frac{\lambda}{\sigma^2} \right)$, $c_n = -\tau \beta f$.

Чтобы получить восстановленное изображение, нам нужно найти положительное решение уравнения (5) следующим выражением:

$$u^{(n+1)} = \frac{-b_n + \sqrt{\Delta_n}}{2a_n}, \quad (6)$$

где $\Delta_n = b_n^2 - 4a_n c_n$.

Ясно что, в (5): $a_n > 0$, $c_n < 0$ ($a_n c_n < 0$), поэтому мы всегда получим положительное решение (6).

3. Эксперименты

Чтобы сравнить эффективность алгоритмов, мы используем пикового отношения сигнал/шум (Peak Signal-to-Noise Ratio, PSNR) [7]. В качестве исходного изображения использовалось изображение 'cameraman' размером 256×256 [6]. Все эксперименты были реализованы в Matlab. Мы добавляли шум Пуассона с пиковой интенсивностью I_{\max} в диапазоне от 120 до 30, а затем гауссовский шум со стандартным отклонением $\sigma = \frac{I_{\max}}{10}$. Мы устанавливали $\varepsilon = 1$, шаг времени $\tau = 0,1$ для всех тестовых изображений. Сравнительные значения PSNR рассмотренных методов показаны в таблице 1. Обратим внимание, что все результаты PSNR (в дБ), указанные в таблице 1, были усреднены по десяти реализациям.

Таблица 1

Сравнение значений PSNR методов

I_{\max}	120	60	30	10
σ	12	6	3	1
[1]	25.97	25.44	24.61	22.78
[2]	26.02	25.47	24.76	22.89
[3]	25.99	25.54	24.61	22.86
[4]	26.04	25.63	24.75	22.99
Наш метод	26.21	25.77	25.15	23.18

4. Заключение

Мы обобщили предложенные ранее оптимизационные методы для модели гауссовского шумоподавления и модели пуассоновской шумоподавления на случай смешанной модели гауссовско-пуассоновского шумоподавления. Предложенный метод подходит для обработки изображений, содержащих смешанный гауссовско-пуассоновский шум.

Благодарности

Работа выполнена при поддержке проекта «5-100» НИУ ВШЭ.

Литература

1. Rudin L., Osher S., and Fatemi E. Nonlinear total variation-based noise removal algorithms // Physica D. — 1992. — Vol. 60, — Pp. 259–268.
2. Le T., Chartrand R. and Asaki T.J. A variational approach to reconstructing images corrupted by Poisson noise // J. Math. Imaging. Vis. — 2007. — Vol. 27. — Pp. 257–263.
3. Wang W., He C. A Fast and Effective Algorithm for a Poisson Denoising Model With Total Variation // J. IEEE Signal Processing Letters. — 2017. — Vol. 24, issue 3. — Pp. 269–273.

4. Dang T., Dvoenko C. Image noise removal based on total variation // Computer Optics. — 2015. — Vol. 39, issue 4. — Pp. 564–571.
5. Immerker J. Fast noise variance estimation // Computer Vision Image Understand. — 1996. — Vol. 64 (2). — Pp. 300–302.
6. Image Databases [Электронный ресурс], URL: <http://www.imageprocessingplace.com/>, (дата обращения: 20.02.2017).
7. Bovik A.C., Wang Z. Modern Image Quality Assessment // Synthesis Lectures on Image, Video, and Multimedia Processing. — Morgan and Claypool Publishers, 2006. — 156 p.

UDC 004.4

A Computationally Efficient Algorithm for Image Restoration Using Total Variation

Cong Thang Pham*, A. Kopylov[†], Thi Thu Thao Tran[†]

** Faculty of Computer Science*

National Research University Higher School Economics, Moscow, HSE

*† Institute Applied Mathematics and Computer Science
Tula State University*

Email: pacotha@gmail.com

In this paper, we present an effective algorithm for image denoising using total variation. The proposed scheme allows to efficiently remove Poisson noise as well as Gaussian noise simultaneously, with the help of a new kind of a data-fidelity term, that is suitable for Gaussian-Poisson noise. Experimental results show the efficiency of our proposed algorithm.

Key words and phrases: Image denoising, Gaussian noise, Poisson noise, Total variation, Gradient flow.

UDC 004.8

On the Safe Achieving the Required Formation Shape by Multiple UAVs (Planar Case)

M. V. Khachumov, V. M. Khachumov

*Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: khachumov_mv@rudn.university, khachumov_vm@rudn.university

The problem of forming a desired system in a group of UAVs is formulated and considered. The general solution scheme of the formulated problem is proposed. Basic solution of the problem is obtained by solving the task of optimal assignment of UAVs to goal positions in the desired formation. Collision avoidance can be ensured by sequential modification of the basic plan based on special rules.

Key words and phrases: unmanned aerial vehicles, formation, group, collision avoidance, assignment problem.

1. Introduction

The required planar or spatial deployment of multiple unmanned aerial vehicles (UAVs) is called a UAV formation. In practice, when forming the desired topology of UAVs, symmetrical geometric shapes (column, rhombus, wedge, circle, sphere, etc.) are often used. The task of specifying the structure in a group of UAVs comes down to calculating the desired goal positions or giving a distance matrix. Hence, certain complex problems may be posed, that are related to the issue of uniform or quasi-uniform allocation of UAVs [1].

In the paper, we assume that a multi-UAV formation is given and the mission for a group of homogeneous arbitrarily located UAVs is to reach safely their goal positions. Such a task is called obtaining of a given shape and includes at least two steps. At first, the problem is to optimally assign UAVs to available positions in the formation. The solution depends on choosing general cost matrix, which estimates UAVs costs for reaching goal positions. At the second step it is necessary to plan flight vehicles trajectories guaranteeing safety. We assume that all UAVs must take required positions in the formation for the same time. Collision avoidance can be ensured by constructing for each vehicle of a tube determined by a circular object. For the simplest rectilinear motions collision detection implies determination of convergence points. When a possible collision is found, the task of resolving the conflict situation comes up. One can swap the assigned goals for conflicting pairs of UAVs on condition of insignificant reduction of the optimization quality. Other way of convergence avoidance is to control UAVs velocities by using special rules.

2. Formulation of the problem and the general solution scheme

Let n identical UAVs $P = \{p_1, p_2, \dots, p_n\}$ be arbitrarily located in some limited region in the plane. The state of UAV p_i at time instant t is described by the following variables: coordinates $(x_i(t), y_i(t))$, speed $v_i(t)$ and the angle of direction $\alpha_i(t)$. Let us consider the following dynamic system

$$x_i(t), y_i(t), \dot{x}_i(t), \dot{y}_i(t), \alpha_i(t), \quad i = 1, \dots, n. \quad (1)$$

We assume that goal positions $S = \{s_1(x_1^s, y_1^s), \dots, s_n(x_n^s, y_n^s)\}$ in the desired formation are given. The problem is to transit the system (1) from the state $x_i(0), y_i(0), \dot{x}_i(0), \dot{y}_i(0), \alpha_i(0)$ that determines the initial location of the vehicles, to the final one

$x_i(T), y_i(T), \dot{x}_i(T), \dot{y}_i(T), \alpha_i(T)$ for a given time T , while minimizing the length of the total traveled path: $\sum_{i=1}^n L_i = ((x_i(0) - x_i(T))^2 + (y_i(0) - y_i(T))^2)$, so that $(x_i(T), y_i(T)) \in S, (x_i(T), y_i(T)) \neq (x_j(T), y_j(T)), i, j = 1, \dots, n$.

Let us give a mathematical description for the simplified UAV motion in a nonperturbed environment: $x_i(t) = v_i \cos(\alpha_i)$; $y_i(t) = v_i \sin(\alpha_i)$. We consider 2D geometric model of a UAV as a circle of radius R (with some margin accounting for the UAV geometry). Safety distance $d_{ij}(t)$ between two vehicles will be determined by the value $d_{ij}(t) > 2R$. When impassable obstacles and wind flows are absent, the distance matrix $D(t) = \|d_{ij}(t)\|, i, j = 1, \dots, n$ is symmetric. For detecting collisions, it is necessary to calculate time instants and points of convergence with account for the UAV models.

The general solution scheme of the formulated problem is defined as follows.

1. Give initial and desired UAVs locations.
 2. Solve the assignment problem and determine velocities of UAVs to reach the desired positions. At this step basic solution of the problem is formed.
 3. Determine time instants and points of UAVs convergence. If collisions exist then go to item 4, otherwise go to item 6.
 4. Swap assigned goals for conflicting pairs of UAVs. If collisions still exist then go to item 5, otherwise go to item 6.
 5. Apply certain rules to determine UAVs velocities ensuring collision avoidance. Calculate new trajectories to reach the desired positions.
 6. End
- Let us consider the problems of assigning UAVs to available positions in the formation and collision avoidance.

3. Assigning UAVs to goal positions and avoiding collisions

Assume that we know the goal positions in the desired structure and positions of all UAVs. The problem is to optimally assign UAVs to available points in the formation. The problem can be presented in the form of a bipartite graph, where vertices of one layer correspond to UAVs $\{p_1, p_2, \dots, p_n\}$ and vertices of another layer correspond to the positions of the formation $\{q_1, q_2, \dots, q_n\}$. The edges of the graph correspond to the elements of the cost matrix $C = \|c_{ij}\|, i, j = 1, \dots, n$, where c_{ij} correspond to the flight of UAV i to point j . In addition, we assign $c_{ij} = \infty$, if the flight of UAV i to position j is impossible or undesirable. We introduce $x_{ij}(i, j = 1, 2, \dots, n)$, such that $x_{ij} = 1$ if UAV i is assigned to point j , and $x_{ij} = 0$ otherwise. A typical assignment task is formulated as a linear programming problem [2]:

$$c = \sum_{i=1}^n \sum_{j=1}^n c_{ij} x_{ij} \rightarrow \min, \quad (2)$$

where: $\sum_{i=1}^n x_{ij} = \sum_{j=1}^n x_{ij} = 1, \forall i, j = 1, 2, \dots, n$.

To solve the problem (2) the Hungarian method can be used [3]. As a result of solving the problem a table is obtained with marked zeros that correspond to the assigned items.

When UAVs are assigned to their goal positions, motions trajectories become known. We assume that all UAVs should arrive at their destinations at the same time T . Collisions between UAVs are potentially possible, when their distances decrease in time. Condition of the collision between two UAVs of a circular shape is represented in a form of the inequality with respect to t

$$\sqrt{(x_i^0 + v_i t \cos \alpha_i - x_j^0 - v_j t \cos \alpha_j)^2 + (y_i^0 + v_i t \sin \alpha_i - y_j^0 - v_j t \sin \alpha_j)^2} \leq 2R. \quad (3)$$

A conflict situation exists if the inequality (3) has real roots in the interval $[0, T]$. In that case decisions are made by using special rules.

Rule 1. Improve the basic plan by trial permutation of the assigned goals for conflicting pairs of UAVs. This rule should reduce the number of collisions.

Rule 2. When vehicles are converging at the same altitude, the vehicle to the other's right has the right-of-way. UAV that must give way reduces the speed to Δv . One should find this value so that the inequality (3) has no real roots in $[0, T]$. Decrease in speed will lead the UAV to take the desired position later than expected. After avoiding a collision the UAV should increase the speed to compensate that.

Rule 3. UAV that gives way, should make turn to the right if the trajectories intersect at an obtuse angle, or to the left if the trajectories intersect at an acute angle. UAV that has a right-of-way should keep its velocity. It is proposed to deflect vehicle's direction by a minimal value $\Delta\alpha$ so that the inequality (3) has no real roots in $[0, T]$. After avoiding a collision the UAV should correct its velocity to reach the goal position in time T . In that case, UAV route is composed of several rectilinear sections.

At first *Rule 1* is applied. If collisions are absent other rules are skipped.

4. Conclusions

An approach to safe flying of a group of vehicles to the desired positions under restriction on travel time is considered. This approach implies obtaining the basic solution of the problem and its sequential modification aimed at reducing the risk of collision. The improvement of the solution is performed by various ways that include reassigning of goals and controlling velocities.

Acknowledgments

The work is supported by RFBR grants No 18-07-00025, No 17-29-07003 and by "RUDN University Program 5-100".

References

1. D. Ya. Ivanov, Methods for constructing spatial formations in groups of quadcopters, Southern Federal University, 2016.
2. H. Kuhn, The hungarian method for the assignment problem, J. Nav. Res. Logist. (2) 83–97, 1955.
3. M. V. Khachumov, N. S. Abramov, D. A. Makarov, Controlling flight vehicle spatial motion along a given route, J. Autom. Remote Control. 76 (6) 1070–1080.

УДК 004.8

Задача безопасного формирования заданного строя БПЛА (плоский случай)

М. В. Хачумов, В. М. Хачумов

*Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: khachumov_mv@rudn.university, khachumov_vm@rudn.university

Сформулирована и рассмотрена задача формирования заданного строя в группе БПЛА. Предложена общая схема решения сформулированной задачи. Опорный план решения формируется путем оптимального назначения БПЛА на заданные точки формации. Безопасность группового полета может быть обеспечена путем последовательной модификации опорного плана с применением специальных правил.

Ключевые слова: БПЛА, строй, группа, безопасность, задача о назначениях.

УДК 519.673

Визуализация потока жидкости через тонкостенный сосуд с переменным сечением

Р. Н. Храбров, М. Б. Фомин

*Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: Shaman1234@gmail.com

В работе рассмотрены алгоритмы численной визуализации потока вязкой жидкости через тонкостенный сосуд с переменным сечением.

Ключевые слова: rendering, pipe, variable cross-section.

1. Введение

В процессе разработки медицинских программных средств может возникнуть задача нахождения времени течения лекарства от места введения к месту назначения. Помимо этого, возможна и задача визуализации этого процесса, т.е. для заданной системы необходимо найти координаты частиц лекарства в каждый момент времени t , после чего отобразить частицы на устройстве вывода изображения.

Особенностью движения жидкости через такой сосуд является то, что из-за его упругих стенок непрерывно меняется диаметр сечения. Данный факт осложняет решение задачи.

2. Постановка физической задачи

Примечание: для решения подобной задачи удобнее всего использовать цилиндрические координаты, где ось z будет направлена вдоль сосуда, а координаты r и θ будут являться полярными координатами, описывающими сечение.

Рассмотрим трёхмерную систему уравнений Навье–Стокса в цилиндрических координатах:

$$\left\{ \begin{array}{l} \rho \left(\frac{\partial v_r}{\partial t} + v_r \frac{\partial v_r}{\partial r} + \frac{v_\theta}{r} \frac{\partial v_r}{\partial \theta} + v_z \frac{\partial v_r}{\partial z} - \frac{v_\theta^2}{r} \right) = \\ \quad = \rho g_r - \frac{\partial p}{\partial r} + \mu \left[\frac{1}{r} \frac{\partial}{\partial r} (r \frac{\partial v_r}{\partial r}) + \frac{1}{r^2} \frac{\partial^2 v_r}{\partial \theta^2} + \frac{\partial^2 v_r}{\partial z^2} - \frac{2}{r^2} \frac{\partial v_\theta}{\partial \theta} - \frac{v_r}{r^2} \right], \\ \rho \left(\frac{\partial v_\theta}{\partial t} + v_r \frac{\partial v_\theta}{\partial r} + \frac{v_\theta}{r} \frac{\partial v_\theta}{\partial \theta} + v_z \frac{\partial v_\theta}{\partial z} + \frac{v_r v_\theta}{r} \right) = \\ \quad = \rho g_\theta - \frac{1}{r} \frac{\partial p}{\partial \theta} + \mu \left[\frac{1}{r} \frac{\partial}{\partial r} (r \frac{\partial v_\theta}{\partial r}) + \frac{1}{r^2} \frac{\partial^2 v_\theta}{\partial \theta^2} + \frac{\partial^2 v_\theta}{\partial z^2} - \frac{2}{r^2} \frac{\partial v_r}{\partial \theta} - \frac{v_\theta}{r^2} \right], \\ \rho \left(\frac{\partial v_z}{\partial t} + v_r \frac{\partial v_z}{\partial r} + \frac{v_\theta}{r} \frac{\partial v_z}{\partial \theta} + v_z \frac{\partial v_z}{\partial z} \right) = \\ \quad = \rho g_z - \frac{\partial p}{\partial z} + \mu \left[\frac{1}{r} \frac{\partial}{\partial r} (r \frac{\partial v_z}{\partial r}) + \frac{1}{r^2} \frac{\partial^2 v_z}{\partial \theta^2} + \frac{\partial^2 v_z}{\partial z^2} \right], \\ \quad \frac{1}{r} \frac{\partial (r v_r)}{\partial r} + \frac{1}{r} \frac{\partial (r v_\theta)}{\partial \theta} + \frac{\partial v_z}{\partial z} = 0. \end{array} \right.$$

В данной системе уравнений:

t – прошедшее время;

ρ – плотность жидкости в данной точке;

μ – коэффициент динамической вязкости в данной точке;

$p = (p_r, p_\theta, p_z)$ – вектор, задающий давление, оказываемое на материальную точку;

$v = (v_r, v_\theta, v_z)$ – вектор скорости жидкости в данной точке;

$g = (g_r, g_\theta, g_z)$ – ускорение, полученное от внешних сил (в частном случае – ускорение свободного падения).

Помимо системы уравнений, необходимы граничные условия. Как было сказано ранее, в данной задаче граничные условия переменны и являются функцией не только от r , θ и z , но и от времени t . Тем не менее, задачу можно упростить, рассмотрев только осесимметричный поток, тем самым избавившись от зависимости от координаты θ . Тогда граничные условия примут следующий вид:

$$F(t, r, z) \equiv r - R(t, z) = 0.$$

Теперь можно сформулировать задачу: для каждого момента $t > 0$ найти координаты частиц жидкости в потоке, протекающем согласно уравнениям 2 с граничными условиями 2.

3. Постановка информационной модели

Рассмотрим конечную цель. Требуется на основе начальных данных построить изображение потока жидкости в сосуде с упругими стенками. Исходя из этого, можно разделить решение задачи на две части — численное решение физической модели и построение изображения на основе полученных данных.

Рассмотрим первую часть. При численном решении подобной задачи обычно разбивают объём жидкости на набор частиц, каждая из которых имеет определённые координаты и несёт информацию о свойствах среды в окрестности данных координат. К этим свойствам можно отнести скорость потока, давление, вязкость и т.д.

Помимо этого, если в сцене присутствует сосуд, то необходима информация о его сечении (граничных условиях). Поскольку, как было сказано выше, рассматривается осесимметричный поток, то эту информацию можно свести к дискретному набору значений радиуса r для дискретного набора значений продольно-осевой координаты z .

Таким образом, можно сформулировать первую часть задачи: для данного набора частиц при заданных параметрах системы и граничных условиях в момент времени t найти их координаты, свойства среды и граничные условия в момент времени $t + \Delta t$.

Теперь рассмотрим вторую часть задачи. Здесь информация о свойствах среды больше не требуется и всё, что необходимо на этом этапе — координаты частиц для визуализации самого потока и информация о граничных условиях для визуализации стенок сосуда.

Вторая часть задачи формулируется следующим образом: для полученного набора координат частиц и информации о граничных условиях получить изображение на устройстве вывода.

Объединив две задачи, получаем одну общую: для данного набора частиц при заданных параметрах системы и граничных условиях в момент времени t получить изображение сцены на устройстве вывода для момента времени $t + \Delta t$. При этом частота смены изображений должна быть приемлемой (минимум 30 изображений в секунду, т.е. время, затраченное на решение данной задачи, не должно превышать 33.3 мс).

4. Алгоритм

Для первой части поставленной задачи использовался модифицированный движок PhysX, в котором есть реализации алгоритмов численного решения задачи моделирования потока жидкости. Эти алгоритмы необходимо было модифицировать для случая переменных граничных условий. Преимуществом такого подхода является то, что все вычисления обрабатываются на графическом процессоре, а значит, позволяют решить проблему одновременной обработки большого числа частиц.

Для второй части задачи использовался собственный алгоритм, реализованный поверх средства визуализации трёхмерной графики Direct3D 11. Сам алгоритм разделяется на две части: 1) Отрисовка объёма жидкости на основе информации о координатах частиц; 2) Отрисовка стенок сосуда на основе переменных граничных условий.

5. Заключение

Была рассмотрена задача визуализации потока жидкости через тонкостенный сосуд с упругими стенками. Для решения был предложен алгоритм решения с использованием движка PhysX и алгоритм отрисовки на основе библиотеки Direct3D 11. Решение данной задачи можно использовать для визуализации в медицинском ПО.

UDC 519.673

Rendering of the Fluid Flow Through a Thin Walled Pipe with Variable Cross-Section

R. N. Khrabrov, M. B. Fomin

*Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: Shaman1234@gmail.com

The algorithms of digital rendering of the viscous fluid flow through a thin walled pipe with variable cross-section were presented in this work.

УДК 004.89

Моделирование естественно-языковой коммуникации с помощью глубокого обучения

Е. В. Чистова

*Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: elenachistov@gmail.com

Построение естественноязыковой диалоговой системы, способной поддержать связный диалог с пользователем, является актуальной проблемой искусственного интеллекта. Подобные системы в зависимости от назначения разделяют на два типа: диалоговые агенты общего назначения, поддерживающие разговор на общие темы, и задачеориентированные диалоговые агенты. В данной работе описываются приложения глубокого обучения к проблеме создания моделей общетематического диалога. Описаны подходы к моделированию диалога, особое внимание уделено применению к задаче моделей нейронных сетей, получивших большое распространение среди исследователей в последние годы. Приведены общие сведения о применении нейросетевой модели "кодировщик-декодировщик" описана её архитектура. Проведены эксперименты на русском языке. Указаны пути дальнейшего развития системы.

Ключевые слова: диалоговые системы, обработка естественного языка, генерация текста, нейронные сети, искусственный интеллект.

1. Введение

В последние годы открытых диалоговых ресурсов стало достаточно для разработок моделей, основанных на больших данных. Многие исследователи используют данные из социальных сетей для построения статистических диалоговых моделей [1]. Другие, как [2], обращают внимание на более современные техники, предполагающие использование векторных представлений слов для численного определения их семантической и синтаксической близости.

В данной работе рассматривается подход, подразумевающий генерацию оригинальных последовательностей слов. При таком подходе зачастую используют методы, разработанные изначально для машинного перевода [1]. В частности, в этой работе особое внимание уделяется архитектуре глубокого обучения типа «кодировщик-декодировщик» [3]. Для обучения таких моделей не требуется предварительная разметка текстов, модель обучается на вопросно-ответных символических последовательностях и после обучения порождает ответы на основе реплик пользователя. Модели после небольшой модификации можно подавать на вход и дополнительные данные о собеседнике (возраст, пол, профессия), если таковые имеются. Так как модель оперирует векторами, она не ориентирована на конкретный язык.

Для автоматической оценки генеративных диалоговых моделей чаще всего прибегают либо к методам, разработанным для задач машинного перевода (BLEU и его модификации, Avg, перплексия). Многие исследователи критично относятся к использованию метрик качества машинного перевода при оценке диалоговых систем. Так, все подобные метрики, использованные в исследовании [4], обнаружили либо слабую корреляцию между метриками для машинного перевода и экспертными оценками, либо её полное отсутствие; в [5] также указывают на неэффективность применения метрик BLEU и перплексии, ввиду их нечувствительности к контексту.

2. Обучение модели «кодировщик-декодировщик»

Seq2seq — это архитектура нейронной сети, состоящей из двух рекуррентных сетей (RNN). Сеть-кодировщик (encoder) принимает на вход последовательность и далее на каждом шаге продуцирует по одному символу (слову). Её задача — сформировать из последовательности символов вектор признаков фиксированной размерности, который содержит только важную информацию из этой последовательности (context vector). Вектор подается декодировщику (decoder), который на каждом шаге генерирует символ (слово) в зависимости от контекстного вектора. Использование RNN для кодировщика и декодировщика позволяет оперировать последовательностями слов произвольной длины [6].

Для экспериментов с русским языком были использованы данные из открытой базы субтитров (100000 реплик) и комментарии русскоязычного информационного сообщества Пикабу (800000 реплик). Была использована готовая реализация¹ модели seq2seq, описанной в [2]. Были обучены пять моделей, результаты представлены в табл. 1, примеры диалоговых обменов представлены в табл. 2. Как уже было отмечено, в настоящее время не существует специальных метрик для оценки качества диалоговых систем. Используемый рутинный метод оценки состоял в создании списка из двухсот общих вопросов, взятых из обычного диалога в мессенджере и подсчёта доли уместных ответов модели. Модели 1, 2 были обучены на небольшом корпусе русскоязычных субтитров; для обучения первой модели выбран также небольшой размер словаря (число знакомых модели слов). Это обусловлено наличием скромной вычислительной мощности. В архитектуру моделей, выделенных жирным, добавлен механизм внимания (attention mechanism). Эти модели показывают предсказуемо лучший результат. Модели, не использовавшие предобученные векторы² слов русского языка, показали пренебрежимые результаты.

Таблица 1

Сравнение seq2seq моделей

Номер	Данные	Размер словаря	Предобученные векторы	Уместность
1	OpenSubs	15000	webriuwikiruscorpога	6%
2	OpenSubs	30000	webriuwikiruscorpога	18%
3	Пикабу	30000	webriuwikiruscorpога	13%
4	Пикабу	25000	araneum skipgram	30%
5	Пикабу	50000	araneum skipgram	38%

Таблица 2

Примеры работы обученных seq2seq моделей

Номер	Примеры диалоговых обменов
3	- Какого цвета желтая машина? - Красного - Дай сюда! - Держи.
4	- Сколько будет два плюс два? - Четыре - Какого цвета небо? - Bloodborne

¹<https://github.com/Conchylicultor/DeepQA>

²<http://rusvectors.org/ru/models/>

3. Заключение

В данной работе были рассмотрены различные подходы к построению диалогового агента, в частности, применение рекуррентной нейронной сети. Были проведены эксперименты с моделированием диалога на общие темы с использованием русскоязычных открытых данных, результаты которых позволяют заключить, что при работе с нейронными архитектурами, оперирующими векторными представлениями символьных последовательностей, целесообразно использовать предобученные векторных представления для конкретного языка, при этом на качество обучения значительно влияет размер словаря.

В будущих работах планируется дальнейший поиск пригодной для задачи диалогового моделирования качественной метрики, а также эксперименты с комбинированием различных подходов к моделированию диалога в рамках одного диалогового агента.

Литература

1. Ritter A., Cherry C., Dolan W. B. Data-driven response generation in social media // Proceedings of the conference on empirical methods in natural language processing. — Association for Computational Linguistics, 2011. — Pp. 583–593.
2. Sordani A. et al. A neural network approach to context-sensitive generation of conversational responses // In Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT 2015).
3. Sutskever I., Vinyals O., Le Q. V. Sequence to sequence learning with neural networks // Advances in neural information processing systems. — 2014. — Pp. 3104–3112.
4. Li J. et al. A diversity-promoting objective function for neural conversation models // In Proc. of NAACL-HLT 2016.
5. Li J. et al. Deep reinforcement learning for dialogue generation // In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing.
6. Shang, L., Lu, Z., and Li, H. Neural responding machine for short-text conversation. In Proceedings of ACL, 2015

UDC 004.89

Modeling Natural Language Communication with Deep Learning

E. V. Chistova

*Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: elenachistov@gmail.com

Building natural language dialogue systems that can converse coherently with user is an actual problem of artificial intelligence. Such systems generally fall into two main categories: open-domain dialogue agents that allow humans to ask general questions and closed-domain agents. In this paper the various deep learning applications to the problem of building open-domain dialogue agent are described. In particular, I focus on systems with neural networks, which have become very popular among various researchers in recent years. It provides general information about encoder-decoder recurrent neural network architecture. Experiments with Russian language are described and the ways of further development of system are suggested.

Key words and phrases: dialogue systems, natural language processing, natural language generation, neural networks, artificial intelligent.

UDC 004

Full Text Patent Classification

V. V. Yadrintsev*[†], I. V. Sochenkov[†]

* *Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

[†] *Institute of System Analysis
Federal Research Center "Computer Science and Control" of RAS
60-letiya Oktyabrya pr. 9, Moscow, Russia, 117312*

Email: vvyadrincev@gmail.com, sochenkov@isa.ru

In this paper, the problem of patent documents classification is considered on the basis of a extended by nominal subgroups vector representation of full-text documents. The classification process begins by extracting keywords and phrases from the documents using by means of automatic text processing. Significant of keywords and phrases are determining according to statistical measure. The topical similarity of documents based on vectors with keywords and phrases is estimating. In this work, the three lowest levels of international patent classification are used as a set of classes

Key words and phrases: patent classification, IPC, topical similarity evaluation.

1. Introduction

Due to the rapid development of information technology and the increase in the number of electronic documents, such as patents, it is impossible to manually process and manage large amounts of text documents. The automated assignment of classes to patents is an important issue, especially for patent offices.

Consequently, initial researches in this direction was conducted by the patent offices [1–3]. The classification methodology with the extraction of keywords from patent documents using a neural back propagation network is known [4]. Good results were obtained, but only few documents were involved in the evaluation of the classifier. In work [5], a study was carried out on the patent classification using a linguistic classification system; different sections (annotation, description, metadata) were used as texts; triplets and references were taken into account. The study was conducted within the framework of the CLEF-IP 2011 competition [6]; EPO and WIPO collections were used as patents. The study [5] showed that the use of full descriptions of patent documents improves the quality of classification: they reached an F1-score of 0.7 for the subclass level in their best run. In [7], a study was conducted on the topic of measurement of patent similarity using keyword comparisons (titles and annotations) of patents.

The purpose of this work is to investigate the use of vectorized document presentation extended by nominal groups for the classification of patent documents at several levels of the international patent classification in open collections of Russian Federal Institute of Industrial Property.

2. Structure of the International Patent Classification

The International Patent Classification (IPC) is a hierarchical patent classification system used in over 100 countries to classify the content of patents in a uniform manner. IPC covers all areas of knowledge whose objects can be protected by protection documents. In the IPC there are five basic levels of the hierarchy. In this paper we take into account 3,4,5 levels of IPC: subclasses, main groups and subgroups.

The number of classes in the sample of documents does not correspond to the number in the used hierarchy of the IPC. This is due to the fact that not for each class there are documents representatives. If there are 641 subclasses in the used hierarchy, then 636 in the sample; in the hierarchy — 6,581 main groups, and in the sample — 5,843; in the hierarchy — 52 544 classes, and in the sample - only 16 765 classes.

3. New approach for classification

To create training data, a method of searching for topically similar documents is used, which is based on a comparison of vectors of lexical descriptors [8]. Let K be the number of topically similar documents. We are considering K from 1 to 10. In this paper, the following ways of forming training data are considered:

1. The best similarity sum for each label (*sum_best*);
2. The best similarity for each label (*best_of_label*);
3. The amount of votes for each label (*count of votes*);
4. The similarity sum for each label (*sum of sims*).

4. Classification evaluation

We are considering the following most popular methods: K -nearest neighbors (KNN), linear support vector (LSVC), logistic regression (LogReg), gradient boosting (GB), random forest (RF). For the experiments we are using open libraries, such as sklearn and lightGBM.

Table 1 shows the main results of the experiments. The first column indicates the classification method, the number of using random documents (in thousands) and level of the hierarchy of the IPC. Columns 2,3,4,5 - types of the data for training, where values in columns: the number of the most topically similar (K) documents (in parentheses) and $F1$ -measure with macro averaging.

Table 1
Results, $F1$ -measure

$F1$ (macro)	<i>sum_best</i>	<i>best_of_label</i>	<i>count of votes</i>	<i>sum of sims</i>
KNN, 50, subclass	(6) 0.6507	(1) 0.6817	(1) 0.6817	(4) 0.6881
RF, 50, subclass	(4) 0.5242	(2) 0.5767	(1) 0.6427	(7) 0.6248
LSVC, 50, subclass	(5) 0.6324	(1) 0.621	(1) 0.6817	(5) 0.6637
LogReg, 50, subclass	(9) 0.6014	(2) 0.4918	(4) 0.6645	(9) 0.6377
GB, 50, subclass	-	-	-	(4) 0.6792
KNN, 10, maingroup	(8) 0.5691	(2) 0.5952	(1) 0.5876	(4) 0.6016
KNN, 10, subgroup	(8) 0.4757	(6) 0.4873	(1) 0.4681	(5) 0.4892

5. Conclusions

The vector representation of patent documents extended by nominal groups allows to obtain a comparable quality of classification with other methods. If the input of data is the assessment of topical similarity of documents, the method of K -nearest neighbors is the most appropriate method of classification.

Future work in the continuation of research on this work: measuring other metrics for the quality of classification; carrying out experiments on other collections.

Acknowledgments

The publication has been prepared with the support of the “RUDN University Program 5-100” and funded by RFBR according to the research projects №16-29-12929 of_m.

References

1. Krier M., Zacca F. Automatic Categorisation Applications at the European Patent Office. — 2002.
2. Fall C., Benzineb K. Literature survey: Issues to be Considered in the Automatic Classification of Patents. — 2002.
3. Fall C. J., Töröcsvári A., Benzineb K., Karetka G. Automated Categorization in the International Patent Classification // SIGIR Forum. — 2003. — Apr. — Vol. 37, no. 1. — P. 10–25.
4. Trappey A. J., Hsu F.-C., Trappey C. V., Lin C.-I. Development of a patent document classification and search platform using a back-propagation network // Expert Systems with Applications. — 2006. — Vol. 31, no. 4. — P. 755 – 765. — Computer Supported Cooperative Work in Design and Manufacturing.
5. Verberne S., D'hondt E. Patent Classification Experiments with the Linguistic Classification System LCS in CLEF-IP 2011. — 2011. — 01.
6. Piroi F. CLEF-IP 2010: Retrieval Experiments in the Intellectual Property Domain. — 2010. — 10.
7. Arts S., Cassiman B., Gomez J. C. Text matching to measure patent similarity // Strategic Management Journal. — 2018. — Vol. 39, no. 1. — P. 62–84.
8. Sochenkov I. V., Zubarev D. V., Tihomirov I. A. Exploratory patent search // Informatics and Applications. — 2018. — Vol. 1. — P. (in print).

УДК 004

Полнотекстовый классификатор патентных документов

В. В. Ядринцев*†, И. В. Соченков†

** Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

*† Институт системного анализа
Федеральный исследовательский центр «Информатика и управление» РАН
пр-т 60-летия Октября, 9, г. Москва, Россия, 117312*

Email: vvyadrincev@gmail.com, sochenkov@isa.ru

В настоящей статье рассматривается задача классификации патентных документов на основе расширенного именованными подгруппами векторного представления полнотекстовых документов. Процесс классификации начинается с извлечения ключевых слов и словосочетаний из документа с помощью автоматической обработки текста, затем выявляются значимые ключевые слова и словосочетания на основе статистической меры, далее оценивается тематическая близость документов на основе векторов с ключевыми словами и словосочетаниями. В настоящей работе в качестве набора классов используются три самых низких уровня международной патентной классификации.

Ключевые слова: классификация патентов, МПК, оценка тематической близости.

УДК 004.9

Проектирование базы данных для системы автоматизации процесса голосования

Ю. Д. Янко

*Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: yanko.julia@yandex.ru

Описывается подход к проектированию базы данных для голосования в организациях, показана ER-модель базы данных, рассмотрена функциональность, поддерживающая работу с базой данных.

Ключевые слова: поддержка принятия решений, процесс голосования, ER-модель, база данных, Delphi, Paradox 7.

1. Введение

На сегодняшний день, самым распространенным способом коллективного принятия управленческих решений для организации является голосование. Само голосование может служить инструментом или способом принятия решения. Существует немало средств, позволяющих провести голосование, однако вопрос о создании удобного и хорошо адаптированного программного обеспечения для коммерческой организации остается открытым. Разработка системы автоматизации процесса голосования, а также ее реализация для частных компаний позволит сократить время принятия решений и повысит интеллектуальные свойства корпоративной информационной системы [1]. Такая система помимо автоматизации процесса обработки и подсчета голосов будет хранить данные, на основе которых можно строить статистику и анализировать ситуацию для принятия в будущем оптимальных решений [2]. В статье рассмотрена структура базы данных для голосования, а также функциональность, реализованная в среде Delphi, работающая с информацией базы данных.

2. Основная часть

2.1. Проектирование базы данных для системы автоматизации процесса голосования

Для голосования, как универсального способа и инструмента системы поддержки и принятия коллективных управленческих решений спроектирована схема базы данных. Она может быть легко встроена в любую из систем голосования как инструмента и как способа поддержки и принятия коллективных решений. Голосование можно разделить на несколько этапов: формирование списка избирателей, создание бюллетеня, процесс выбора альтернатив голосующими, создание протокола. Именно реализация каждого из этапов находит отражение в структуре базы данных, представленной в виде ERM-модели. В модели был учтен тот факт, что бюллетень формируется на основании вопросов, выдвинутых на голосование и альтернатив. Протокол же создается на основе ответов каждого из пользователей. Все ответы необходимо сохранять, а затем производить подсчет голосов. Также был учтен тот факт, что идентификационный номер каждого собрания должен совпадать с идентификационным номером бюллетеня и протокола, поскольку они являются обязательными инструментами голосования [3].

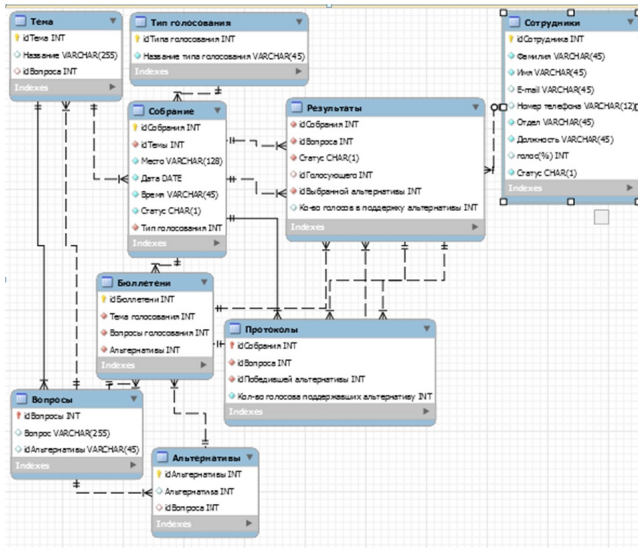


Рис. 1. Структура базы данных для системы голосования [5]

Таблицы, описывающие хранение информации:

- *Таблица Сотрудники* хранит всю информацию о сотрудниках компании: ФИО, email, номер телефона, а также указание их должностей, отдел, в котором работает каждый, и статус. Если под должностью понимается штатная единица организации, выполняющая ряд соответствующих обязанностей, то под статусом понимается комплекс должностных полномочий — общий для всех должностей [4]. Именно по статусу будут определяться участники голосования.
- *Таблица Собрание* хранит всю информацию о собрании, на котором проводится голосование: место, время, дата, статус людей, которые могут принять участие в голосовании, и самое главное – тема собрания (поскольку на одном собрании могут обсуждаться несколько тем).
- *Таблица Бюллетени* содержит информацию о вопросах для голосования и предоставленных альтернатив по каждому вопросу. Вопросов поднятых на голосовании также может быть несколько.
- *Таблица Протоколы* хранит информацию об уже обработанных результатах голосования. Информацию о победившей альтернативе и количестве голосов ее поддержавших.
- *Таблица Тип голосования* предназначена для хранения типов голосования. Поскольку для голосования в коммерческих организациях используется определенные типы, то эта таблица будет содержать всего четыре записи: простое голосование, вековое голосование, кумулятивное голосование, голосование с блокированием.
- *Таблица Тема* хранит название темы голосования и номера вопросов, рассматриваемые в рамках темы.
- *Таблица Вопросы* предназначена для хранения вопросов и номера альтернатив по каждому из них.

- *Таблица Альтернативы* хранит варианты ответов для вопросов. Одна запись может служить альтернативой для нескольких вопросов.
- Таблица, описывающая процесс голосования:
- *Таблица Результаты* содержит информацию о самом собрании, содержит вопросы, выдвинутые на собрании, статус и идентификационные номера голосующих, голос, отданный в поддержку определенной альтернативы конкретного голосующего.

2.2. Функциональность, реализованная в среде Delphi, работающая с информацией базы данных

Процесс подсчета и обработки голосов, с учетом типа голосования, будет реализован функционалом системы (в рамках программного обеспечения). Стоит отметить, что созданная модель базы данных подходит для открытого, очно-заочного голосования, поскольку администратор программы видит ответы каждого из голосующих.

Для связи и работы с базой данных в среде Delphi используется выпускаемая компанией Corel реляционная СУБД Paradox 7 [6]. Для таблиц, предназначенных для хранения информации были созданы отдельные классы, благодаря которым пользователь может редактировать, добавлять, удалять записи. В процессе обработки голосов программа обращается к таблице Результаты. В зависимости от типа голосования производится расчет голосов для каждого голосующего, все голоса, отданные в поддержку определенной альтернативы, суммируются, после чего происходит сравнение количества голосов. Альтернатива, набравшая наибольшее количество голосов сохраняется в таблице Протоколы. Программа реализована таким образом, что значения из таблиц Бюллетени и Протоколы могут быть вставлены в шаблон протокола и бюллетеня и представлены пользователю в текстовом формате.

3. Заключение

Спроектированная база данных является универсальной базой данных, которая может быть реализована для любого приложения: web-голосования, мобильного приложения, прикладной программой, и т. д. На данном этапе база данных была реализована в рамках пилотного проекта для голосования компании в клиент-серверном приложении.

Система поддерживает голосование в любой предметной области, где решение принимается коллективным способом.

Литература

1. Новикова Г. М. Основы разработки корпоративных инфокоммуникационных систем. — 2008, РУДН, 152 с.
2. Новикова Г. М. Руководителю нужны интеллектуальные системы // Нефтегазовая вертикаль. — 2005. — №3. — С. 64–66.
3. Федеральный закон «Об основных гарантиях избирательных прав и права на участие в референдуме граждан Российской Федерации» от 12.06.2002 N 67-ФЗ (последняя редакция) [Электронный ресурс]. — Режим доступа: http://www.consultant.ru/document/cons_doc_LAW_37119/, свободный.
4. Классификация должностей гражданской службы (категории и группы) [Электронный ресурс]. — Режим доступа: <http://studbooks.net/>, свободный.
5. Инструментарий ARIS [Электронный ресурс]. — Режим доступа: http://businessanalysis.osngrad.info/images/BA_images/aris_methods.pdf, свободный.
6. Учебное пособие «Основы проектирования реляционных баз данных». [Электронный ресурс]. — Режим доступа: <http://www.delphiplus.org/>, свободный.

UDC 004.9

Database Design for Automation of the Voting Process in Application Software

J. D. Yanko

*Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: yanko.julia@yandex.ru

The approach to designing a voting database in organizations is described, the ER model of the database is shown, the functionality supporting the work with the database is considered.

Key words and phrases: Decision Support System, the voting process, ER-model, database, Delphi, Paradox 7.

**Распределённые, высокопроизводительные
вычисления и аналитика больших данных**

**Distributed, high-performance computing technologies
and big data analytics**

УДК 004.4

Виртуальные рабочие столы гетерогенного кластера HybriLIT

Г. Адам*[†], Д. В. Беляков*, М. Валя*[‡], П. В Гончаров[§], П. В. Зрелов*,
В. В. Кореньков*, М. А. Матвеев*,
Д. В. Подгайный*, О. И. Стрельцова*

** Лаборатория информационных технологий
Объединённый институт ядерных исследований
ул. Жолио-Кюри, д. 6, Дубна, Московская область, Россия, 141980*

† IFIN-HH, Бухарест-Мэгуре, Румыния

‡ Институт экспериментальной физики САН в Кошице, Словакия

*§ Гомельский государственный технический университет им. П. О. Сухого,
Республика Беларусь*

Email: adamg@jinr.ru, dmitry@jinr.ru, mvala@saske.sk, kaliistrogoblin3@gmail.com, zrelov@jinr.ru,
korenkov@jinr.ru, matveevma@jinr.ru, podgainy@jinr.ru, strel@jinr.ru

Пользователи, проводящие расчёты на высокопроизводительном кластере, сталкиваются с проблемой визуализации большого объёма полученных данных и задержками при работе в интерактивном режиме с результатами счёта. Аппаратные характеристики и программное обеспечение вычислительных систем не всегда позволяют обеспечить необходимый уровень производительности при визуализации данных и комфортной работы в интерактивном режиме. Для решения этой проблемы был применён подход, позволяющий использовать персональный виртуальный рабочий стол для работы с графическими данными.

В данной работе рассматриваются различные реализации виртуальных рабочих столов и используемые технологии удаленного доступа для взаимодействия клиентского компьютера с вычислительным сервером.

Ключевые слова: высокопроизводительный кластер, виртуальный рабочий стол.

1. Введение

При выполнении расчётов пользователю часто приходится оперировать большими объёмами как входных, так и выходных данных. При этом он сталкивается с проблемой недостаточной производительности одного вычислительного узла, связанной с аппаратными ограничениями (число ядер процессора, объём оперативной памяти, скорость доступа к системе хранения данных). Как правило, данную проблему можно решить с помощью использования нескольких вычислительных узлов с программным обеспечением, задействующим технологию параллельной обработки данных. В качестве результатов счёта пользователь получает файлы с выходными данными, размеры которых могут достигать нескольких сотен гигабайт или же нескольких терабайт. Получение результатов счёта для последующего анализа или визуализации может занять длительное время, либо создать трудности с хранением большого объёма данных на компьютере пользователя. По этим причинам для дальнейшей работы или визуализации выходных данных используется удалённый доступ к виртуальному рабочему столу.

Поскольку работа пользователя часто носит итерационный характер, рост объёма данных приводит к критическому увеличению сетевого трафика, а время передачи данных между вычислительным кластером и компьютером пользователя становится сопоставимым со временем выполнения расчётов.

Использование удалённого доступа к виртуальным рабочим столам (Virtual Desktop Infrastructure, VDI) позволяет использовать ресурсы гетерогенного кластера HybriLIT [1] для выполнения расчётов и локальные ресурсы виртуальной машины, в том числе графический ускоритель NVIDIA Tesla M60 [2], для дальнейшей работы и визуализации выходных данных.

2. Реализации VDI

Гипервизор — программное обеспечение, позволяющее одновременное выполнение одной или нескольких операционных систем на одном вычислительном сервере.

Рендеринг — процесс построения изображения на основе цифровых данных. Разделяют программный и аппаратный рендеринг, в первом случае формирование изображения выполняет центральный процессор (CPU), во втором случае формирование изображения выполняет графический процессор (GPU).

На гетерогенном кластере HybriLIT первой реализацией VDI была система, использующая гипервизор 2-го типа — KVM, установленная на вычислительные серверы без графических процессоров. Компьютер пользователя получал удалённый доступ к виртуальному рабочему столу по протоколу SPICE [3]. Переход на протокол X2GO [4] позволил повысить уровень комфортной работы в среде VDI. Однако недостатком такой реализации является отсутствие графического процессора на вычислительных серверах, что существенно снижает скорость рендеринга интерфейса виртуального рабочего стола.

Второй реализацией VDI стала система, построенная на гипервизоре 1-го типа — XEN, задействованном в сочетании целого набора программного обеспечения компаний Citrix [5], NVIDIA и Microsoft [6]. Система была развернута на вычислительном сервере, содержащем графические ускорители NVIDIA Tesla M60.

Для подготовки виртуальных машин использовался Citrix XenCenter, для управления инфраструктурой Citrix XenDesktop, для разделения ресурсов графических ускорителей между виртуальными машинами NVIDIA Quadro профили, для аутентификации пользователей и разделения прав доступа Microsoft Domain Controller. Компьютер пользователя получает доступ к виртуальному рабочему столу с помощью клиента Citrix XenReceiver по протоколу Citrix HDX 3D Pro.

Схематичное представление реализаций VDI приведено на рисунке 1.

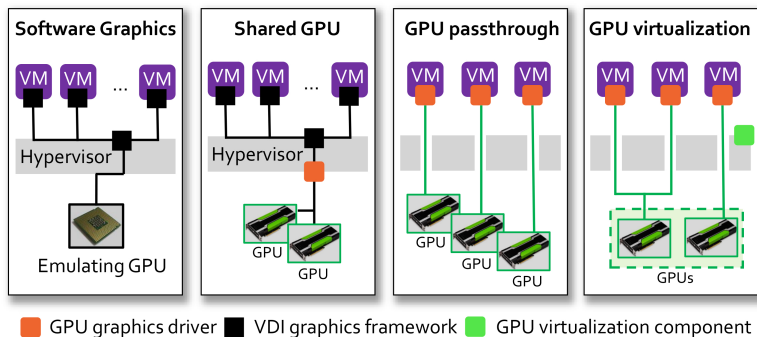


Рис. 1. Схематичное представление реализаций VDI с использованием центрального процессора и графического ускорителя

В 2018 году введена в эксплуатацию новая компонента гетерогенного кластера HybriLIT — система виртуальных рабочих столов на основе гипервизора 1-го типа XEN, существенно расширяющая функциональные возможности вычислительного кластера. Использование VDI позволяет пользователям получать удалённый доступ к прикладному программному обеспечению для параллельной обработки данных с использованием ресурсов гетерогенного кластера HybriLIT и для комфортной работы с графическими приложениями.

Внедрение и отладка системы виртуальных рабочих мест проводилась на реальных задачах применения машинного обучения для генетической классификации и трекинга для установки VM@N [7, 8].

3. Заключение

Добавление новой компоненты в структуру гетерогенного кластера HybriLIT позволило удалённо работать с прикладным программным обеспечением, таким как Comsol [9], Maple [10], Mathematica [11], Matlab [12] и выполнять параллельные расчёты с использованием ресурсов гетерогенного кластера HybriLIT.

Литература

1. Сайт HybriLIT. URL: <http://hybrilit.jinr.ru>
2. Сайт компании NVIDIA. URL: <http://nvidia.com>
3. Протокол SPICE. URL: <https://www.spice-space.org>
4. Протокол X2GO. URL: <https://wiki.x2go.org>
5. Сайт компании Citrix. URL: <https://www.citrix.ru>
6. Сайт компании Microsoft. URL: <https://www.microsoft.com>
7. Two-stage approach to image classification by deep neural networks, EPJ Web of Conferences 173, 01009 (2018) <https://doi.org/10.1051/epjconf/201817301009>
8. Baranov D. et. al., Novel approach to the particle track reconstruction based on deep learning methods, Proceedings of the XXVI International Symposium on Nuclear Electronics & Computing (NEC 2017), Pp. 37–45. URL: <http://ceur-ws.org/Vol-2023/37-45-paper-6.pdf>.
9. Пакет COMSOL. URL: <https://www.comsol.ru>
10. Пакет Maple. URL: <https://www.maplesoft.com>
11. Пакет Mathematica. URL: <https://www.wolfram.com>
12. Пакет Matlab. URL: <https://www.mathworks.com>

UDC 004.4

Virtual Desktops of the Heterogeneous Cluster HybriLIT

Gh. Adam^{*†}, D. V. Belyakov^{*}, M. Vala^{*‡}, P. V. Goncharov[§], P. V. Zrellov^{*},
V. V. Korenkov^{*}, M. A. Matveyev^{*}, D. V. Podgainy^{*}, O. I. Streltsova^{*}

^{*} *Laboratory of Information Technologies
Joint Institute for Nuclear Research*

6 Joliot-Curie str., Dubna, Moscow region, Russian Federation, 141980

[†] *IFIN-HH, Bucharest-Magurele, Romania*

[‡] *Institute for Theoretical and Experimental Physics, Kosice, Slovakia*

[§] *Sukhoi State Technical University of Gomel, Republic of Belarus*

Email: adamg@jinr.ru, dmitry@jinr.ru, mvala@saske.sk, kaliostrogoblin3@gmail.com, zrellov@jinr.ru,
korenkov@jinr.ru, matveevma@jinr.ru, podgainy@jinr.ru, strel@jinr.ru

Users who perform calculations on a high-performance cluster deal with the problem of visualizing a large amount of data received, and delays when working in interactive mode with the results of calculations. Hardware and software specifications of computer systems do not always allow to provide the necessary level of performance when visualizing data and comfortable work in interactive mode. To solve this problem, an approach on using a personal virtual desktop to work with graphics data has been adopted providing

This paper examines different virtual desktop implementations and remote access technologies used to communicate between the client computer and the computing server.

Key words and phrases: high-performance cluster, virtual desktop.

УДК 519-642

Параллельная реализация модели микроскопического оптического ядро-ядерного потенциала на основе технологий MPI и OPENMP

М. В. Башашина*[†], Е. В. Земляная*[†], К. В. Лукьянов*[†]

* Лаборатория информационных технологий
Объединённый институт ядерных исследований
ул. Жолио-Кюри, д. 6, Дубна, Московская область, Россия, 141980

[†] Институт системного анализа и управления
Государственный университет «Дубна»
ул. Университетская, д. 19, Дубна, Московская область Россия, 141982

Email: bashashinmv@jinr.ru, elena@jinr.ru, luku@jinr.ru

Метод двойного фолдинга для построения микроскопического оптического потенциала ядро-ядерного рассеяния является одним из популярных методов расчета этой важнейшей ядерно-физической характеристики. Расчет прямого потенциала не представляет вычислительных трудностей, так как он сводится к вычислению интегрального выражения. Обменный потенциал строится на основании численного решения нелинейного интегрального уравнения, что требует при массовых расчетах существенных затрат компьютерного времени. Разработанная параллельная вычислительная схема реализована на основе технологии MPI (Message Passing Interface) и OpenMP (Open Multi-Processing). Эффективность разработанных реализаций подтверждена методическими расчетами на многопроцессорном кластере HybriLIT (ЛИТ ОИЯИ, Дубна).

Ключевые слова: ядро-ядерное рассеяние, оптический потенциал, итерационный процесс, параллельные вычисления.

1. Введение

В данной работе представлена параллельная реализация одного из активно используемых методов расчета микроскопического оптического ядро-ядерного потенциала, выполненная в двух вариантах, на основе двух популярных технологий параллельного программирования — MPI и OpenMP. Кратко описывается метод расчета потенциала на основе модели двойного фолдинга и приводятся результаты методических расчетов, демонстрирующие эффективность MPI- и OpenMP-версий.

2. Математическая модель

Потенциал модели двойного фолдинга V_{DF} строится как сумма изоскалярной и изовекторной составляющих, каждая из которых включает прямую V_D и обменную V части [1–4]. Изоскалярный потенциал имеет вид:

$$V_{DF}(r) = V_D(r) + V_{EX}(r),$$

где

$$V_{DF}(r) = \int d^3r_p d^3r_t \{ \rho_p(\vec{r}_p) \rho_t(\vec{r}_t) \nu_{NN}^D \},$$

$$V_{EX}(r) = \int d^3r_p d^3r_t \{ \rho_p(\vec{r}_p, \vec{r}_p + \vec{s}) \rho_t(\vec{r}_t, \vec{r}_p - \vec{s}) \nu_{NN}^{EX} \times \exp[i\vec{K}(r) \vec{s} / M] \},$$

где $M = A_p A_t / (A_p + A_t)$. Каждая функция плотности налетающего ядра ρ_p и ядра-мишени ρ_t с атомными массами A_p и A_t есть сумма нейтронной и протонной

плотности (в выражении для V^{EX} — матрицы плотностей). $\vec{K}(r, V^F)$ — локальный импульс относительного движения

$$K^2(r) = \frac{2m_0M}{\eta} [E - V_{DF}(E, r) - U_C(r)],$$

где U_C — кулоновский потенциал. Эффективные нуклон-нуклонные потенциалы ν_{NN}^D и ν_{NN}^{EX} зависят от энергии и плотности сталкивающихся объектов. Изовекторная часть потенциала получается в формулах для V_D и V заменой суммы нейтронной и протонной плотностей на их разность, а также другими выражениями для эффективных NN -потенциалов. Поскольку, локальный импульс ядро-ядерного движения K зависит от потенциала V_{DF} , выражение для обменного потенциала представляет собой нелинейное интегральное уравнение, которое схематично можно записать в форме

$$V_{EX}(r) = \int_0^\infty db [\Phi(V_{EX}(b), r)],$$

где Φ в свою очередь представляет собой интегральное выражение кратности 3. Для численного решения уравнения вводится равномерная сетка по координате. Численное интегрирование осуществляется с помощью формулы Симпсона. Решение уравнения основано на методе простых итераций. В качестве начального приближения V_{EX}^0 выбирается заранее рассчитанный прямой потенциал V_D , вычисление которого сводится к однократному расчету интегрального выражения. На каждой итерации с номером k рассчитывается правая часть уравнения и, тем самым, получается $(k + 1)$ -е приближение:

$$V_{EX}^{k+1}(r) = \int_0^\infty db [\Phi(V_{EX}^k(b), r)].$$

Итерации завершаются, когда для заранее заданного малого положительного числа ε выполнено условие

$$\max_r |V_{EX}^{k+1} - V_{EX}^k| < \varepsilon.$$

Вычисления показывают, что при выборе $V_{EX}^0 = V_D$ и $\varepsilon = 0.0001$ достаточно 6–8 итераций. Естественно, вместо сингулярного интервала интегрирования $[0, \infty)$ расчеты проводятся на конечном интервале $[0, R_{\max}]$, где R_{\max} — достаточно большое число, обеспечивающее правильное асимптотическое поведение МДФ-потенциала и его близость к 0 в области R_{\max} . Сложность расчета возрастает с ростом числа узлов дискретной сетки по координате пропорционально степени 4, т.е. при увеличении вдвое числа узлов дискретной сетки время счета возрастает примерно в 16 раз.

3. Численные результаты

Эффективность реализации описанного в [5] параллельного алгоритма подтверждается расчетами на кластере HUBRILIT Многофункционального информационно-вычислительного комплекса Лаборатории информационных технологий Объединенного института ядерных исследований (Дубна). Подробная информация о кластере доступна на сайте <http://hubrilit.ru>.

На рис. 1 представлена зависимость ускорения вычислений от количества задействованных параллельных узлов, а также полученный МДФ-потенциал для случая упругого рассеяния экзотического ядра ${}^6\text{He}$ с кинетической энергией 50 МэВ/нуклон на ядре-мишени ${}^{28}\text{Si}$. В качестве плотности $\rho_{p,t}$ ядра ${}^6\text{He}$ использована линейная комбинация функций гауссовского типа.

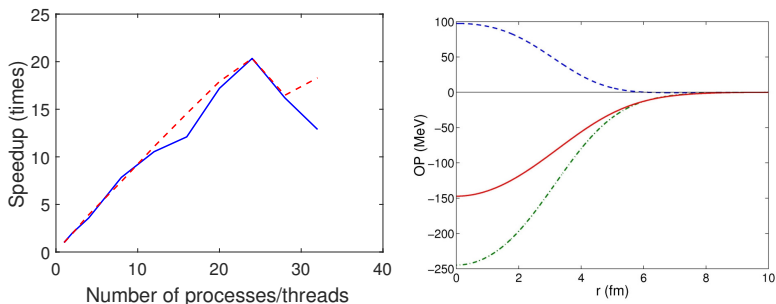


Рис. 1. Слева: ускорение вычислений при расчете МДФ-потенциала на кластере HybridLIT для случая ${}^6\text{He}+{}^{28}\text{Si}$ при энергии 50 МэВ/нуклон с числом узлов дискретной сетки $N = 601$ (сплошная кривая — MPI, штриховая — OpenMP). Справа: рассчитанный МДФ-потенциал (штриховая кривая — прямая часть V^D , штрихпунктирная — обменная часть V^{EX} , сплошная — результирующий потенциал V^F)

Для плотности ядра-мишени ${}^{28}\text{Si}$ использовалась ферми-функция. Расчеты проводились на интервале интегрирования $[0, 10]$ с числом узлов равномерной дискретной сетки $N=601$. Результаты расчетов на основе MPI- и OpenMP-версий совпадают между собой и с результатами работы исходной последовательной компьютерной программы DFROT, что подтверждает корректность созданных MPI(C++) и OpenMP(C++) программ. Расчеты проводились с использованием Intel-компилятора версии 2017.2.050.

4. Заключение

В работе на основе проведенных методических расчетов показано, что разработанные параллельные версии метода расчета микроскопического оптического потенциала упругого рассеяния ядер на ядрах в рамках модели МДФ на основе технологий параллельного программирования MPI и OpenMP являются эффективными и обеспечивают значительное ускорение вычислений при работе в параллельном режиме. В настоящее время ведутся работы по подготовке разработанных программ к размещению в библиотеке JINRLIB.

Благодарности

Работа выполнена при финансовой поддержке гранта РФФИ (международный проект № 17-52-18057-Болг_а).

Литература

1. Кхоа Д. Т. и Князьков О. М. Обменные эффекты в ядро-ядерных потенциалах и ядерное раджное рассеяние // Физика элементарных частиц и атомного ядра. — Т. 21. — 1990. — С. 1456.
2. Khoa D. T., Satchler G. R. Generalized folding model for elastic and inelastic nucleus–nucleus scattering using realistic density dependent nucleon–nucleon interaction // Nucl. Phys. A. — 2000. — Vol. 668. — Pp. 3–41.

3. Земляная Е. В., Лукьянов В. К., Лукьянов К. В. Ядро-ядерное рассеяние и оптический потенциал фолдинга // Ядерная физика. — 2006. — Т. 69, вып. 2. — С. 262–275.
4. Лукьянов К. В. Модель ядро-ядерного потенциала двойного фолдинга: основные формулы, итерационный метод и программа вычисления // Сообщение ОИЯИ Р11-2007-38, Дубна, 2007.
5. Земляная Е. В., Лукьянов К. В., Башахин М. В. MPI-реализация расчета микроскопического оптического ядро-ядерного потенциала в рамках модели двойного фолдинга // Системный анализ в науке и образовании. — 2016. — Вып. 3. — С. 41–48.

UDC 519-642

Parallel Implementation of the Microscopic Optical Nucleus-Nucleus Potential Model on the Basis of MPI and OPENMP Technologies

M. V. Bashashin^{*†}, E. V. Zemlyanaya^{*†}, K. V. Lukyanov^{*†}

** Laboratory of Information Technologies
Joint Institute for Nuclear Research*

6 Joliot-Curie str., Dubna, Moscow region, 141980, Russian Federation

*† System Analysis and Management Institute
Dubna State University*

19 Universitetskaya str., Dubna, Moscow region, 141982, Russian Federation

Email: bashashinmv@jinr.ru, elena@jinr.ru, luky@jinr.ru

The double folding model is the well-known method for constructing the microscopic optical potential of nucleus-nucleus scattering. The calculation of this characteristic consists of two stages. The direct potential is obtained by means of an integral expression. The exchange potential calculation is based on numerical solution of the nonlinear integral equation which requires significant computer costs in mass calculations. The developed parallel computing scheme is implemented on the basis of two technologies: MPI (Message Passing Interface) and OpenMP (Open Multi-Processing). The effectiveness of the developed implementations is confirmed by methodical calculations on the multiprocessor cluster HybriLIT (LIT JINR, Dubna).

Key words and phrases: nucleus-nuclear scattering, optical potential, iterative process, parallel computations.

УДК 519.711.3

Оценка адекватности модели и объекта исследования

В. Н. Добрынин, А. А. Миловидова, И. А. Соколов

*Кафедра системного анализа и управления
Государственный университет «Дубна»
ул. Университетская, д. 19, Дубна, Московская область, Россия, 141980*

Email: arbatsolo@yandex.ru, milanna@uni-dubna.ru, isokolov@jinr.ru

Проблема оценки адекватности модели реальным процессам, явлениям, и т.д. является главной при исследовании этих объектов на основе моделирования. Корректность аргументации (доказательности) результатов задач исследования на основе моделирования базируется на структурно-функциональной адекватности имитационной модели объекту и адекватности состояния объекта и состояния модели на заданном интервале времени. При этом, поскольку модель, как правило, характеризуется системой параметров и допустимой областью их значений, что означает наличие в пространстве параметров таких их значений, при которых состояние модели совпадает с наблюдаемым состоянием объекта на всём заданном интервале времени при заданных начальных условиях. В случае отсутствия таких значений параметров необходимо осуществить коррекцию модели. В работе предлагается решение поставленной проблемы.

Ключевые слова: адекватность, модель, объект, оценка, мера сходства.

1. Введение

Решение задач анализа сложных объектов (таких например как технологические процессы производств, природные катастрофы, транспортные системы, социальные явления и т.д.) как правило, осуществляется на основе моделирования. При этом, задача адекватности структурно-функциональной организации имитационной модели и объекта исследования в нашем случае считается разрешённой. Следует отметить, что создание модели осуществляется при наличии не только объекта исследования, а также проблемы, связанной с объектом. На ментальной карте (рис. 1) представлены необходимые компоненты и связи между ними для реализации технологии оценки адекватности модели объекту исследования. На карте представлена ситуация, когда модель не является адекватной объекту. В этом случае возможны различные варианты, представленные на карте. Сходство состояния модели и наблюдаемого состояния объекта на заданном интервале времени определяется некоторым порогом, ниже которого модель и объект считаются адекватными, а выше — не адекватными. При этом мера вычисления сходства может быть чёткой, либо нечёткой. В нашем случае рассматривается чёткая мера сходства.

2. Основная часть

2.1. Формальная постановка задачи

1. Задано описание конечного класса (K) процессов (P_i), $P_i \in K$, $i = 1, 2, \dots, n$.
2. Задано описание класса (Ω) моделей процессов (M_i), $M_i \in K$, $i = 1, 2, \dots, n$.
3. Задано описание соответствия между объектом и моделью $S(P_i, M_i)$.
4. Задано описание общего пространства состояния (U) объекта и модели.
5. Определено текущее состояние объекта исследования
 $\bar{Y}^*(t) = (Y_1^*(t), Y_2^*(t), \dots, Y_m^*(t))$ ($P \in K$) при определённых условиях R для $t \in [t_0, T]$.
6. Задано $\bar{Y}^*(t_0) = \bar{Y}_0^* \in D$ (D — множество допустимых начальных значений).
7. Определен вектор параметр $\bar{B} = (b_1, b_2, \dots, b_i, \dots, b_m)$.
8. R определяет допустимые значения параметров в виде системы ограничений

$$R = \begin{cases} b_1^{\min} \leq b_1 \leq b_1^{\max} \\ \dots \\ b_i^{\min} \leq b_i \leq b_i^{\max} \\ \dots \\ b_m^{\min} \leq b_m \leq b_m^{\max} \end{cases}$$

9. При произвольных начальных значениях вектора состояния и вектора параметра имеем множество всевозможных состояний объекта т.е. $\bar{Y}^*(t, \bar{Y}_0^* \in D, \bar{B} \in R)$.
10. Определено состояния модели объекта $\bar{Y}(t), t \in [0, T], (M \in \Omega)$ при определённых условиях R' для $t \in [t_0, T]$.
11. Задано $\bar{Y}(t_0) = \bar{Y}_0 \in D'$ (D' – множество допустимых начальных значений).
12. Определён вектор параметр модели $\bar{A} = (a_1, a_2, \dots, a_i, \dots, a_r)$.
13. R' определяет допустимые значения параметров в виде системы ограничений

$$R' = \begin{cases} a_1^{\min} \leq a_1 \leq a_1^{\max} \\ \dots \\ a_i^{\min} \leq a_i \leq a_i^{\max} \\ \dots \\ a_r^{\min} \leq a_r \leq a_r^{\max} \end{cases}$$

14. При произвольных начальных значениях вектора состояния и вектора параметра имеем множество всевозможных состояний модели объекта т.е. $\bar{Y}(t, \bar{Y}_0 \in D, \bar{A} \in R')$.

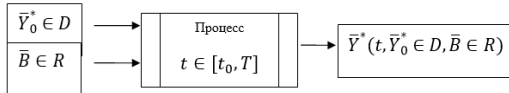


Рис. 1. Схема получения различных состояний объекта

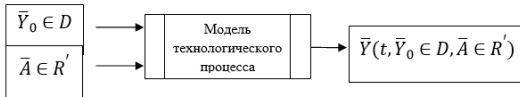


Рис. 2. Схема получения различных состояний модели объекта

Пусть задан некоторый процесс $P \in K$ и модель $M \in \Omega$ этого процесса. Мера сходства μ модели M объекту (обозначим как $\mu(M, P)$) – это процедура, определяющая численную величину, которая характеризует близость модели объекту. Поскольку объект и модель характеризуется своими состояниями, определим меру сходства следующим образом.

Мера сходства состояния модели M и объекта P , для фиксированного вектора $\bar{B}^* = (b_1^*, b_2^*, \dots, b_i^*, \dots, b_m^*)$ – это величина $\mu(\bar{Y}(t, \bar{Y}_0 \in D, \bar{A} \in R'), \bar{Y}^*(t, \bar{Y}_0^* \in D, \bar{B} \in R))$, которая характеризует близость состояния модели объекту P (рис. 3).

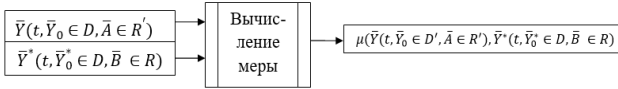


Рис. 3. Схема вычисления меры сходства объекта и модели

Согласно сказанному для определения меры сходства необходимо привести реальное время, на котором задано состояние объекта к модельному.

Для множества R' построим сетку: $S(\Delta_i) : \Delta_i = \frac{a_i^{\max} - a_i^{\min}}{n_i} i = 1, 2, \dots m$.

$u_i - j$ -й узел сетки ($j = 1, 2, \dots n_1, n_2 \dots n_m$).

Для узла u_i определено значение \bar{A} , т.е. $(a_{j1}, a_{j2} \dots a_{jm})$.

Для узла u_i и начального состояния $\bar{Y}_0 \in D'$ определено состояние модели $\bar{Y}(t, \bar{Y}_0 \in D', (a_{j1}, a_{j2} \dots a_{jm}) \in R') = \bar{Y}_j(t, \dots), j = 1, 2, \dots n_1, n_2 \dots n_m$.

Так же для определения меры сходства необходимо задать шаг дискретизации времени $t_i = t_0 + h(i - 1), h = \frac{T}{N}, N$ — число точек дискретизации.

Найдя меру сходства можно получить меру различия. Данная величина вычисляется, как $\gamma = 1 - \mu(M, P)$ В данной работе модель будет считаться адекватной объекту исследования в пространстве состояний, если для заданной некой количественной величины δ , мера различия $\gamma(M, P)$ меньше или равна δ .

2.2. Вычислительный алгоритм

1. В пространстве параметров модели определяется матрица квадратичных невязок W_{rm} для всех пар состояния модели и объекта (см. табл. 1).

Матрица квадратичных невязок

Таблица 1

	$Y_1^*(t), Y_1(t)$...	$Y_k^*(t), Y_k(t)$...	$Y_m^*(t), Y_m(t)$
\bar{A}_1	$\frac{(\sum Y_1^* - Y_1)^2}{N}$...	$\frac{(\sum Y_k^* - Y_k)^2}{N}$...	$\frac{(\sum Y_m^* - Y_m)^2}{N}$
\bar{A}_2	$\frac{(\sum Y_1^* - Y_1)^2}{N}$...	$\frac{(\sum Y_k^* - Y_k)^2}{N}$...	$\frac{(\sum Y_m^* - Y_m)^2}{N}$
...
\bar{A}_i	$\frac{(\sum Y_1^* - Y_1)^2}{N}$...	$\frac{(\sum Y_k^* - Y_k)^2}{N}$...	$\frac{(\sum Y_m^* - Y_m)^2}{N}$
...
\bar{A}_r	$\frac{(\sum Y_1^* - Y_1)^2}{N}$...	$\frac{(\sum Y_k^* - Y_k)^2}{N}$...	$\frac{(\sum Y_m^* - Y_m)^2}{N}$

2. В пространстве параметров модели определяется вектор минимальных квадратичных невязок $\bar{W}_{\min} = (W_{\min 1}, W_{\min 2}, \dots, W_{\min r})$.
3. В пространстве параметров модели формируется нуль единичная матрица отклонений квадратичных невязок от минимальных W'_{rm} .
4. В пространстве параметров модели формируется вектор коэффициентов близости ($\bar{\mu}$) и различия ($\bar{\gamma}$) модели и объекта.
5. В векторе коэффициентов различия определяется наилучший, для этого задается δ . Будем считать, что модель адекватна объекту P если найдётся такое значение вектора $\bar{A} = \bar{A}' \in R'$ для которого выполняется условие $\bar{\gamma} \leq \delta$, для $t \in [0, T]$.

Таблица 2
 Определение вектора минимальных квадратичных невязок

	$Y_1^*(t), Y_1(t)$...	$Y_k^*(t), Y_k(t)$...	$Y_m^*(t), Y_m(t)$
\bar{A}_1	$\frac{(\sum Y_1^* - Y_1)^2}{N}$...	$\frac{(\sum Y_k^* - Y_k)^2}{N}$...	$\frac{(\sum Y_m^* - Y_m)^2}{N}$
...
\bar{A}_r	$\frac{(\sum Y_1^* - Y_1)^2}{N}$...	$\frac{(\sum Y_k^* - Y_k)^2}{N}$...	$\frac{(\sum Y_m^* - Y_m)^2}{N}$
	$\min \frac{(\sum Y_1^* - Y_1)^2}{N}$		$\min \frac{(\sum Y_k^* - Y_k)^2}{N}$		$\min \frac{(\sum Y_m^* - Y_m)^2}{N}$

Таблица 3
 Нуль-единичная матрица отклонений квадратичных невязок от минимальных

	$Y_1^*(t), Y_1(t)$..	$Y_m^*(t), Y_m(t)$
\bar{A}_1	$W'_{11} = \begin{cases} 1, & \frac{(\sum Y_1^* - Y_1)^2}{N} - W_{\min 1} = 0 \\ 0, & \frac{(\sum Y_1^* - Y_1)^2}{N} - W_{\min 1} > 0 \end{cases}$..	$W'_{1m} = \begin{cases} 1, & \frac{(\sum Y_m^* - Y_m)^2}{N} - W_{\min 1} = 0 \\ 0, & \frac{(\sum Y_m^* - Y_m)^2}{N} - W_{\min 1} > 0 \end{cases}$
...
\bar{A}_r	$W'_{r1} = \begin{cases} 1, & \frac{(\sum Y_1^* - Y_1)^2}{N} - W_{\min r} = 0 \\ 0, & \frac{(\sum Y_1^* - Y_1)^2}{N} - W_{\min r} > 0 \end{cases}$..	$W'_{rm} = \begin{cases} 1, & \frac{(\sum Y_m^* - Y_m)^2}{N} - W_{\min r} = 0 \\ 0, & \frac{(\sum Y_m^* - Y_m)^2}{N} - W_{\min r} > 0 \end{cases}$

Таблица 4
 Формирование векторов близости и различия

	$Y_1^*(t), Y_1(t)$...	$Y_m^*(t), Y_m(t)$	$\bar{\mu}(0 \leq \bar{\mu} \leq 1)$	$\bar{\gamma}(0 \leq \bar{\gamma} \leq 1)$
\bar{A}_1	W'_{11}	...	W'_{1m}	$\sum_{i,j=1}^m W'_{1m}/m$	$1 - \sum_{i,j=1}^m W'_{1m}/m$
...
\bar{A}_r	W'_{r1}	...	W'_{rm}	$\sum_{i,j=1}^m W'_{rm}/m$	$1 - \sum_{i,j=1}^m W'_{rm}/m$

3. Заключение

Разработанный алгоритм может быть использоваться для следующего класса задач:

- проверка соответствия объекта с заданным вектором параметров и построенной модели с тем же вектором параметров;
- поиск для объекта исследования вектора параметров модели;
- вектор параметров модели и объекта исследования не совпадают.

Если возникает ситуация, что не находится ни одной меры различия удовлетворяющей δ , то алгоритм необходимо уточнить за счет уменьшения сетки области параметров и /или шага дискретизации интервала времени.

Литература

1. Синельник С. А., Коврижных О. Е. Определение меры сходства объектов в кластерном анализе [Электронный ресурс] // Бизнес конспект. — Электрон. дан. — М., 2015. — URL: <http://www.konspekt.biz/index.php?text=51064> (дата обращения: 01.03.2017).

UDC 519.711.3

The Assessment of Adequacy of Model to the Object of the Research

V. N. Dobrynin, A. A. Milovidova, I. A. Sokolov

*System Analysis and Management Institute
Dubna State University*

19 Universitetskaya str., Dubna, Moscow Region, 141982, Russian Federation

Email: arbatsolo@yandex.ru, milanna@uni-dubna.ru, isokolov@jinr.ru

Assessment of adequacy of model to an object of a research – one of the main objectives modelling various processes and the phenomena: natural, social, technological, scientific, etc. The argument of results of modelling at transfer on a real object, is based on an assessment of adequacy of model to an object. When the received results don't contradict the expected results, due attention isn't paid to a question of adequacy. When the result is unexpected for the researcher, then first of all the attention to adequacy of model of an object. In work authors offer problem definition and an algorithm of the solution of a problem of an assessment of adequacy of model to an object of a research.

Key words and phrases: adequacy, model, object, assessment, similarity measure.

УДК 519.23

Перекрестная энтропия для отбора признаков в задачах анализа данных

Ю. А. Дубнов

Институт системного анализа

*Федеральный исследовательский центр «Информатика и управление» РАН
пр. 60-летия Октября, д. 9, Москва, Россия, 117312*

Факультет компьютерных наук

*Национальный исследовательский университет «Высшая школа экономики»
Кочновский проезд, д. 3, Москва, Россия, 125319*

Email: yury.dubnov@phystech.edu

В работе рассматривается задача понижения размерности, возникающая в задачах анализа данных при наличии обучающих коллекций больших объемов. Приводится алгоритм отбора признаков, базирующийся на вероятностном подходе и понятии перекрестной энтропии (cross entropy). Предложенный алгоритм тестируется на примере типовых задач классификации и сравнивается с наиболее распространенным аналогом в терминах средней точности классификации.

Ключевые слова: понижения размерности, отбор признаков, классификация, перекрестная энтропия.

1. Введение

Основной целью понижения размерности является уменьшение вычислительной сложности алгоритмов машинного обучения при наименьшей возможной потере информации для обучения [1]. Одним из самых распространенных методов понижения размерности является метод главных компонент (Principal Component Analysis), основанный на вычислении направлений наибольшего разброса признаков и проецировании на них облака точек обучающей коллекции [2]. Однако, непосредственный отбор признаков производится, как правило, по правилу объясненной дисперсии, что не отражает специфику распределения значений признаков по классам.

Одним из альтернативных подходов к отбору признаков является использование понятий из теории информации, таких как информационная энтропия, взаимная информация и расстояние Кульбака-Ляйблера. Основанные на этих понятиях методы отбора признаков приведены в работах [3–5]. Информационные критерии отбора позволяют не только уменьшить вычислительную сложность обучения, но и повысить точность благодаря выявлению неинформативных и избыточных признаков.

В данной работе приводится новый алгоритм отбора признаков, основанный на вычислении перекрестной энтропии (cross entropy) для эмпирических распределений вероятности по каждому признаку. Описание алгоритма приведено в разделе 2, а результаты тестирования и сравнения с методом PCA в разделе 3.

2. Перекрестная энтропия и смесь нормальных распределений

Перекрестная энтропия вычисляется для двух распределений вероятности и позволяет оценить их схожесть. Одним из этих распределений положим эмпирическую функцию распределения вероятности для некоторого признака $p(x)$, а другим — некоторое базовое распределение $q(x)$, схожесть с которым повышает полезность признака.

В качестве базового распределения рассмотрим смесь нормальных распределений, состоящую, например, из двух компонент для задачи бинарной классификации:

$$q(x) = \omega_1 \mathcal{N}(x, \mu_1, \sigma_1^2) + \omega_2 \mathcal{N}(x, \mu_2, \sigma_2^2),$$

где веса компонент соответствуют долям классов в обучающей коллекции, т.е. $\omega_1 + \omega_2 = 1$, а параметры компонент вычисляются на основе меток классов c :

$$\mu_i = \mathbb{E}[x|c = i], \quad \sigma_i^2 = D[x|c = i], \quad i = 1, 2.$$

Отметим, что смесь нормальных распределений выбрана для примера с упрощенной реализацией, в действительности в качестве компонент базового распределения имеет смысл рассматривать различные типы распределений, наилучшим образом описывающих структуру значений данного признака: логнормальное, экспоненциальное, гамма и пр.

Таким образом, критерием отбора признаков является максимизация величины перекрестной энтропии между эмпирическим распределением признака и выбранного базового распределения: $CE = H(p, q) = -\sum_x p(x) \log q(x) \rightarrow \max$.

Данная процедура повторяется для каждого признака в изначальном наборе данных и выбираются наиболее значимые признаки в терминах перекрестной энтропии, которые и используются для последующего обучения.

3. Результаты экспериментов

Предложенный алгоритм был протестирован на типовых задачах классификации из открытого репозитория данных для машинного обучения лаборатории KEEL (Knowledge Extraction based on Evolutionary Learning) [6]. В качестве базового классификатора используется классический метод SVM (Support Vector Mashine) с линейной разделяющей поверхностью [7].

Таблица 1

Результаты экспериментов

Назв.	N	N/A	PCA	PCA+CE
hepatitis	80	82.56 ± 2.22	83.18 ± 1.09	87.89 ± 2.14
appendicitis	106	87.67 ± 0.93	87.41 ± 0.64	87.90 ± 0.91
sonar	208	74.53 ± 1.52	77.31 ± 1.29	78.59 ± 0.90
spectfheart	267	79.23 ± 1.39	77.97 ± 1.05	82.24 ± 1.01
haberman	306	72.53 ± 0.49	72.53 ± 0.50	73.53 ± 0.00
ionosphere	351	88.12 ± 0.80	86.30 ± 0.71	88.12 ± 0.80
bands	365	68.64 ± 0.88	63.01 ± 0.00	68.74 ± 0.88
wdbc	569	97.13 ± 0.43	90.78 ± 0.17	96.08 ± 0.43
wisconsin	683	96.72 ± 0.19	96.69 ± 0.19	97.38 ± 0.10
pima	768	77.06 ± 0.34	74.03 ± 0.18	77.04 ± 0.35

Методы сравниваются в терминах средней точности классификации, которая вычисляется по 10 блокам кроссвалидации (10-fold) и усредняется по результату 100 различных разбиений, благодаря чему все полученные результаты являются воспроизводимыми в пределах среднеквадратичного отклонения.

Результаты экспериментов приведены в табл. 1, где для каждого набора данных приведены объем выборки (N), средняя точность классификации без понижения размерности (N/A), с классическим методом PCA по уровню объясненной дисперсии 95% и с отбором проекций по критерию перекрестной энтропии (PCA+CE).

Как видно, в большинстве примеров переход к энтропийному критерию отбора позволяет существенно повысить точность классификации. С другой стороны, в некоторых примерах итоговая точность классификации уступает классификации изначального набора данных, к таким примерам относятся коллекции ionosphere, bands, wdbc и pima. Это связано с неизбежной потерей информации при понижении

размерности, как по методу PCA, так и по энтропийному критерию, но уже с существенно меньшей потерей.

4. Заключение

В работе рассмотрена задача отбора наиболее информативных признаков и приведен алгоритм ее решения, основанный на вычислении параметров смеси нормальных распределений и величины перекрестной энтропии. Предложенный алгоритм позволяет повысить точность классификации благодаря уменьшению информационных потерь при отборе признаков. Главным преимуществом алгоритма является акцент на распределение значений признаков в зависимости от класса, в то время как отбор признаков по методу главных компонент гарантирует лишь наибольший разброс точек в пространстве.

Благодарности

Работа частично поддержана грантом РФФИ № 17-07-00286.

Литература

1. Bishop C. Pattern Recognition and Machine Learning (Information Science and Statistics). — Springer, 2006. — 758 p.
2. Jolliffe I.T. Principal Component Analysis, Series: Springer Series in Statistics, 2nd ed. — Springer, NY, XXIX, 2002. — 487 p.
3. Abellán J., Castellano J. G. Improving the Naive Bayes Classifier via a Quick Variable Selection Method Using Maximum of Entropy // Entropy. — 2017. — Vol. 19, no. 6. — P. 247.
4. Peng H. C., Long F., Ding C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2005. — Vol. 27(8). — P. 1226–1238.
5. Zhang Y., Li S., Wang T., Zhang. Z. Divergence-based feature selection for separate classes // Neurocomputing. — 2013. — Vol. 101. — Pp. 32–42.
6. Alcalá-Fdez J., Fernandez A., Luengo J., Derrac J., Garcia S., Sánchez L., Herrera F. KEEL Data-Mining Software Tool: Data Set Repository, Integration of Algorithms and Experimental Analysis Framework // Journal of Multiple-Valued Logic and Soft Computing. — 2011. — Vol. 17:2-3. — Pp. 255–287.
7. Christianini N., Shawe-Taylor J. An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. — Cambridge, UK: Cambridge University Press, 2000.

UDC 519.23

Cross Entropy for Feature Selection in Data Analysis Problems

Yu. A. Dubnov

*Institute for Systems Analysis (ISA), FRC CSC RAS,
9 pr. 60-letiya Otyabrya, Moscow, 117312, Russian Federation
Faculty of Computer Science, Higher School of Economics (HSE)
3 Kochnovsky Proezd, Moscow, 125319, Russian Federation*

Email: yury.dubnov@phystech.edu

The paper considers the problem of dimensionality reduction in data analysis problems of large training collections. A feature selection algorithm based on the probabilistic approach and the concept of cross entropy is given. The proposed algorithm is tested for typical classification problems and is compared with the most common method in terms of the average classification accuracy.

Key words and phrases: dimensionality reduction, feature selection, classification, cross entropy.

УДК 004.023

HLIT-VDI – новый сервис IT-экосистемы гетерогенного кластера HybriLIT для работы с прикладным программным обеспечением

М. И. Зуев, М. А. Матвеев, Д. В. Подгайный,
О. И. Стрельцова, Ш. Г. Торосян

*Лаборатория информационных технологий
Объединённый институт ядерных исследований
ул. Жолио-Кюри, д. 6, Дубна, Московская область, Россия, 141980*

Email: shushanik@jinr.ru

В программно-информационную среду кластера HybriLIT [1] введен новый сервис – HLIT-VDI, который предназначен для работы с прикладным программным обеспечением использующем развитые графические интерфейсы. Данный сервис позволяет работать с такими пакетами прикладных программ, как Wolfram Mathematica, Maple, MATLAB, Geant4 и др. через удаленный доступ на виртуальные машины (ВМ) на базе кластера HybriLIT. Созданный сервис обеспечивает проведение как небольших расчетов непосредственно на ВМ, так и выполнение ресурсоемких расчетов на вычислительных узлах кластера.

Ключевые слова: гетерогенный кластер, HybriLIT, программно-информационная среда, виртуализация, пакеты прикладных программ.

1. Введение

В рамках программно-информационной среды (IT-экосистемы) гетерогенно-го кластера HybriLIT, предназначенного для проведения массивно-параллельных расчетов с использованием широкого спектра программного обеспечения (ПО), разработан новый сервис – HLIT-VDI (HybriLIT Virtual Desktop Interface), который предоставляет пользователям удаленный доступ на созданные виртуальные рабочие столы.

Технология VDI (Virtual Desktop Infrastructure) – это форма виртуализации вычислительных ресурсов, в которой все элементы рабочего стола пользователя размещены удаленно, а пользователь подключается к своему рабочему столу с клиентского устройства. Благодаря этой технологии пользователи могут получить доступ и работать на кластере не только через терминал с командной строкой, но и запускать и работать с приложениями с поддержкой графического интерфейса (Graphical User Interface – GUI), а также использовать в своих исследованиях сложные математические (MATLAB, Mathematica, Maple) и физические (Geant4, ROOT) пакеты.

При этом ни рабочий стол пользователя, ни приложения, ни данные на устройстве пользователя локально не хранятся.

2. Обзор характеристик и алгоритм работы с виртуальной машиной

В настоящий момент, для пользователей подготовлено 24 ВМ со следующими характеристиками: 20 Гб оперативной памяти; 6 ядер процессора Intel Xeon E5-2680; NFS в качестве домашней директории; операционная система Scientific Linux 7.4.

Все расчеты могут проводиться как в рамках ресурсов виртуальной машины, так и с помощью вывода задачи в общую очередь кластера HybriLIT. Такое разделение обусловлено тем, что для задач, не требующих привлечения большого количества вычислительных ядер достаточно ресурсов виртуальной машины, однако для ряда ресурсоемких задач требуется привлечение вычислительных ресурсов кластера HybriLIT. С другой стороны, использование технологии VDI позволяет не загружать кластер «легкими» задачами.

Доступ к виртуальному рабочему столу открыт каждому пользователю кластера HybriLIT. Для подключения к виртуальной машине (VM) пользователю необходимо установить и настроить программу-клиент – X2Go [2]. Выполнив необходимые шаги по установке и настройке клиента, пользователь попадает на виртуальную машину, где ему доступен весь набор ПО, установленного на кластере HybriLIT. Далее пользователь продолжает работу с необходимыми ему пакетами. Среди таких пакетов прикладных программ, включающих как бесплатное, так и лицензионное ПО, можно выделить следующие математические пакеты: MATLAB, Mathematica, Maple, COMSOL; и также физические пакеты такие как: ROOT, Geant4.

Спектр задач, решаемых с помощью упомянутых программных пакетов довольно широк. Например, задачи квантового компьютеринга успешно решаются с помощью пакета Mathematica. Учитывая, что задачи такого типа довольно ресурсоемки, был разработан специальный скрипт, который позволяет пользователям отправлять задачу напрямую в очередь кластера HybriLIT, что существенно ускоряет вычисления.

3. Выводы

Разработанный сервис предоставил возможность пользователям работать с лицензионным ПО, а также позволил повысить эффективность использования кластера HybriLIT.

Такая организация рабочего процесса позволила пользователям быстрее и эффективнее считать свои задачи с использованием специализированного ПО, в том числе обладающим развитым графическим интерфейсом, а также разгрузила очередь задач на кластере и позволило повысить эффективность использования вычислительных ресурсов кластера.

Также стоит отметить, что теперь пользователям доступны дорогостоящие программные пакеты такие как Wolfram Mathematica, Maple, MATLAB.

Литература

1. Веб-сайт HybriLIT. URL <http://hybrilit.jinr.ru>
2. Веб-сайт X2Go. URL <https://wiki.x2go.org/doku.php>

UDC 004.023

HLIT-VDI – a New Service of the HybriLIT Ecosystem for Work with Applied Software Packages

M. I. Zuev, M. A. Matveev, D. V. Podgainy, O. I. Streltsova,
Sh. G. Torosyan

*Laboratory of Information Technologies
Joint Institute for Nuclear Research*

6 Joliot-Curie str., Dubna, Moscow Region, Russian Federation, 141980

Email: shushanik@jinr.ru

A new service – HLIT-VDI – has been developed for shared use of applied software packages on the HybriLIT cluster using GUI (graphical user interface). By means of this service, it is now possible to work with applied software packages such as Wolfram Mathematica, Maple, MATLAB, Geant4, etc. via remote access to the virtual machines (VM) in the framework of the HybriLIT cluster. The developed service allows carrying out computations in the frames of VMs and massive computations using the resources of the cluster.

Key words and phrases: heterogeneous cluster, HybriLIT, information-software environment, visualization, applied program packages.

УДК 004.75

Интеграция облачных инфраструктур ЛИТ ОИЯИ и Астанинского филиала ИЯФ

Е. М. Мажитова*[†], Н. А. Балашов*, А. В. Баранов*,
Н. А. Кутровский*, Р. Н. Семенов*[‡]

* *Лаборатория информационных технологий
Объединённый институт ядерных исследований
ул. Жолио-Кюри, д. 6, Дубна, Московская область, Россия, 141980*

[†] *Институт ядерной физики
ул. Ибрагимова, д. 1, Алматы, Республика Казахстан, 050032*

[‡] *Российский экономический университет им. Г. В. Плеханова
Стремянный пер., д. 36, Москва, Россия, 117997*

Email: emazhitova@jinr.ru

В данной статье описывается мотивация и ход работ по интеграции облачной инфраструктуры Астанинского филиала Института ядерной физики (ИЯФ) с распределенной облачной инфраструктурой, развернутой на ресурсах Лаборатории информационных технологий Объединенного института ядерных исследований (ЛИТ ОИЯИ) и некоторых организаций из стран-участниц ОИЯИ.

Ключевые слова: облачные вычисления, интеграция, OpenNebula, cloud bursting.

1. Введение

ОИЯИ участвует в большом количестве научно-исследовательских проектов, во многих из которых важным инструментом для получения значимых научных результатов выступают компьютерные инфраструктуры. В этой связи объединение вычислительных ресурсов ОИЯИ и организаций из его стран-участниц представляется важной и актуальной задачей, решение бы которой позволило существенно ускорить проведение научных исследований.

2. Интеграция облачных инфраструктур

Существует несколько возможных вариантов объединения территориально распределенных вычислительных ресурсов в единую платформу. Это может быть, как использование грид-технологий по аналогии с интеграцией ресурсов в глобальных инфраструктурах, например, WLCG [1], так и облачных технологий, например, по т. н. модели «cloud bursting» или, другой возможный вариант, объединение по модели федеративного устройства (в [2] более подробно описаны возможные варианты интеграции облачных инфраструктур).

Для объединения вычислительных ресурсов партнерских организаций стран-участниц ОИЯИ с целью их совместного использования для решения общих задач, а также для распределения пиковых нагрузок между ними, было выбрано объединение по модели «cloud bursting», для чего облачной командой из ОИЯИ был разработан специальный драйвер [2]. Он позволяет интегрировать облако ОИЯИ с облаками партнеров, функционирующих либо на основе платформы OpenNebula [3] (и в этом случае можно включить мониторинг внешних облачных ресурсов в реальном времени), либо на любой другой облачной платформе, поддерживающей Open Cloud Computing Interface (OSCI). В первом случае драйвер для интеграции основан на ruby-реализации интерфейса OCCI (rOCCI) и OpenNebula XML-RPC. Причина одновременного применения двух указанных интерфейсов заключается в том, что OCCI позволяет увеличить повторное использование кода для различных популярных облачных платформ, которые поддерживают данный интерфейс (например, OpenStack), тогда как функции XML-RPC платформы OpenNebula

являются специфическими конкретно для данной платформы и не могут быть использованы для интеграции с другими. В то же время ОССИ не предоставляет функции мониторинга, в отличие от XML-RPC платформы OpenNebula. Было принято решение использовать rOCCI для операций управления и XML-RPC для мониторинга. Схема драйвера показана на рис. 1.

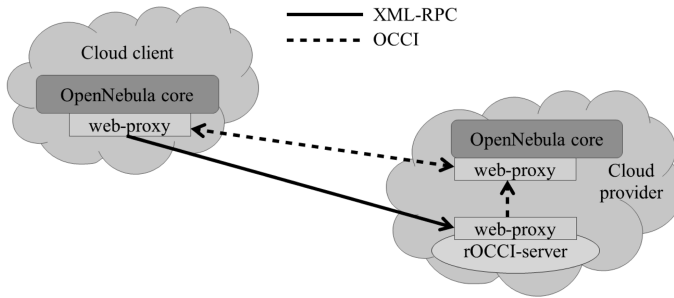


Рис. 1. Схема драйвера

Облако ОИЯИ является ядром распределенной облачной инфраструктуры [4]. На данный момент интеграция облаков организаций из стран-участниц ОИЯИ находится на разных этапах, в частности:

- Институт физики Национальной академии наук Азербайджана (Баку, Азербайджан, интегрирован);
- Институт теоретической физики им. Н. Н. Боголюбова Национальной академии наук Украины (Киев, Украина, интегрирован);
- Российский экономический университет им. Плеханова (Москва, Россия, интегрирован);
- Научно-исследовательский институт ядерных проблем Белорусского государственного университета (Минск, Беларусь, интегрирован);
- Частное учреждение «Nazarbayev University Library and IT services» (Астана, Казахстан, в процессе интеграции);
- Астанинский филиал Института ядерной физики (Астана, Казахстан, в процессе интеграции);
- Институт ядерных исследований и ядерной энергии (София, Болгария, в процессе интеграции);
- Софийский университет «Св. Климент Охридски» (София, Болгария, в процессе интеграции);
- Ереванский физический институт (Ереван, Армения, в процессе интеграции);
- Грузинский технический университет (Тбилиси, Грузия, в процессе интеграции);
- Институт ядерной физики (Ташкент, Узбекистан, переговоры);
- Университет «Св. Климент Охридски» (Битола, Македония, переговоры).

На рис. 2 показаны на карте местоположения участников распределенной облачной инфраструктуры.

В рамках трехлетнего проекта с приоритетным финансированием в период 2017-2019 гг. ведется деятельность, связанная с интеграцией ресурсов ОИЯИ и Астанинского филиала ИЯФ в распределенную облачную инфраструктуру. Ниже представлен план реализации проекта по годам:

1. исследование возможных решений для объединения ресурсов (2017 г.);

2. приобретение оборудования и построение тестового полигона на выбранном решении (2017–2018 гг.);
3. проведение обучающих мероприятий для казахстанских специалистов (2017–2018 гг.);
4. интеграция ресурсов в единую распределенную информационно-вычислительную среду (РИВС) (2018 г.);
5. тестирование РИВС (2018 г.);
6. привлечение потенциальных организаций-участников в проект (ИЯФ Республики Казахстан, ЕНУ им. Л. Н. Гумилева, Государственный университет им. Е. Букегова, Казахский национальный университет им. Аль-Фараби, частные компании) (2018–2019 гг.).

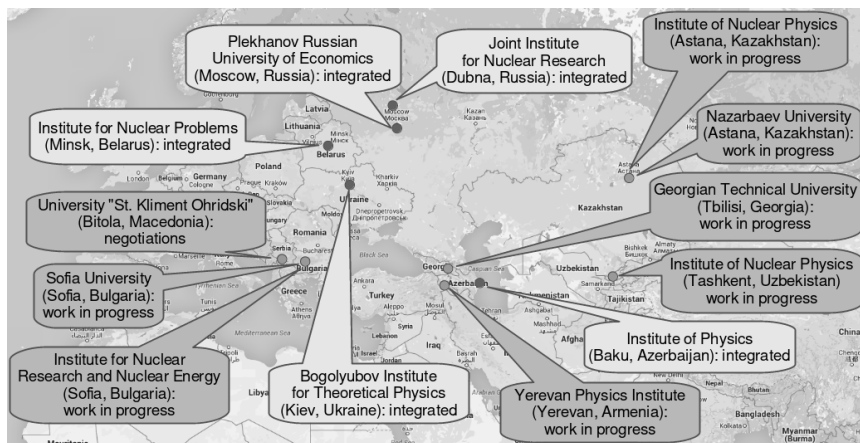


Рис. 2. Карта с участниками распределенной облачной инфраструктуры

3. Заключение

Изучены возможные способы интеграции территориально распределённых вычислительных ресурсов. Выбрано решение на основе объединения облачных инфраструктур по модели «cloud bursting». Был организован и проведен комплекс учебных мероприятий для специалистов из Казахстана по созданию облачных платформ на базе OpenNebula и использованию ресурсов распределённых информационно-вычислительных сред. В качестве прототипа на имеющемся в настоящее время в Астинском филиале ИЯФ оборудовании был развернут испытательный полигон. Интеграция ресурсов будет способствовать ускорению получения новых научных результатов и достижений в различных областях, в которых ведут свою работу ОИЯИ и организации из его государств-участниц, включая Казахстан, а также станет основой для проведения учебных мероприятий для студентов, учёных и ИТ-специалистов по использованию современных информационных технологий для хранения, обработки и анализа больших объемов научных данных.

Литература

1. Bird I. Computing for the Large Hadron Collider // Annual Review of Nuclear and Particle Science. — 2011. — Vol. 61, Pp. 99–118. doi: 10.1146/annurev-nucl-102010-130059.
2. Baranov A. V. et al. Approaches to cloud infrastructures integration // Computer Research and Modeling. — 2016. — Vol. 8 (3). — Pp. 583–590.
3. OpenNebula web-portal. URL: <http://opennebula.org>
4. Baranov A. V. et al. JINR cloud infrastructure evolution // Particles and Nuclei Letters. — 2016. — Vol. 13 (5), Pp. 672–675. doi: 10.1134/S1547477116050071.

UDC 004.75

Integration of the Cloud Infrastructures of the LIT JINR and INP's Astana Branch

E. M. Mazhitova*[†], N. A. Balashov*, A. V. Baranov*, N. A. Kutovskiy*,
R. N. Semenov*[‡]

** Laboratory of Information Technologies
Joint Institute for Nuclear Research
Joliot-Curie 6, Dubna, Moscow region, 141980, Russia*

*† The Institute of Nuclear Physics,
Ibragimova 1, Almaty, 050032, the Republic of Kazakhstan*

*‡ Plekhanov Russian University of Economics,
Stremyanny lane 36, Moscow, 117997, Russia*

Email: emazhitova@jinr.ru

This article describes a motivation and current progress on an integration of the cloud infrastructure of Institute of Nuclear Physics (INP) Astana branch into distributed cloud infrastructure deployed on the basis of the resources of the Laboratory of Information Technologies of the Joint Institute for Nuclear Research (LIT JINR) and some organizations from its Member State.

Key words and phrases: cloud computing, integration, OpenNebula, cloud bursting.

УДК 004.72

Суперкомпьютерное моделирование функционирования беспроводной сенсорной сети в ситуации потери связи с базовой станцией

И. М. Никольский

*Факультет вычислительной математики и кибернетики
Московский государственный университет им. М. В. Ломоносова
Ленинские горы, ГСП-1, Москва, Россия 119991*

Email: oliv_mail@mail.ru

В работе рассматривается цепочка статических сенсоров с несколькими стоками. На режим работы узлов сети и среду передачи данных накладываются условия модели Чиассерини–Гаретто. Учитываются такие особенности функционирования беспроводных сенсорных сетей (БСС) как интерференция радиосигнала, наличие активной и пассивной фаз в работе конечных устройств и т.д. Построена параллельная имитационная модель, предназначенная для оценки параметров работы данной БСС в условиях отсутствия связи с базовой станцией. Модель предназначена для вычислительных систем с распределенной памятью и МРІ-коммуникаций. Тестирование на суперкомпьютере BlueGene/P показывает хорошую масштабируемость модели. Модель позволяет изучить различные характеристики работы БСС (таких как скорость заполнения буферов узлов сети), входящих в системы мониторинга шоссе и других протяженных объектов.

Ключевые слова: беспроводные сенсорные сети, имитационное моделирование, параллельные алгоритмы, суперкомпьютер.

1. Введение

Успехи в области электроники последних лет привели к взрывному росту количества проектов, связанных с автоматизацией различных сторон деятельности человека — от логистики до экологического мониторинга. Основой многих таких проектов являются беспроводные сенсорные сети (БСС). Узлами беспроводных сетей являются так называемые «умные сенсоры» (motes). Эти устройства способны периодически измерять некоторую физическую величину (концентрацию вредных веществ в воздухе, уровень шума и т.д.), производить предварительную обработку данных и передавать информацию на базовую станцию (или в некоторую облачную инфраструктуру), используя другие сенсоры в качестве ретрансляторов.

Как правило, «умный сенсор» состоит из датчика, источника питания, радиомодуля и микроконтроллера. Обмен информации между узлами БСС производится по беспроводному протоколу (такому как Wi-Fi, Стриж, Zwave и т.д.) Наиболее популярным протоколом для построения современных БСС является протокол ZigBee [1].

При проектировании БСС возникают проблемы, отсутствующие в обычных компьютерных сетях: необходимо учитывать ограниченность заряда батареи, наличие препятствий, интерференцию радиосигнала и т.д. Ненадежность каналов связи может привести к отключению сенсорной сети от базового сервера; при этом возможно переполнение буферов устройств, в результате чего произойдет потеря части накопленных данных. Моделированию этой ситуации и посвящена эта работа.

В литературе по БСС (см. [2–5]) принято рассматривать равномерное размещение сенсоров по области обслуживания, имеющей форму круга или прямоугольника. Однако представляют интерес и регулярное размещение устройств, так как при монтаже сенсоры удобнее располагать, следуя некоторой закономерности. В данной работе рассматривается цепочка статических сенсоров. Узлы располагаются вдоль прямой, расстояние между соседними сенсорами одинаково. БСС является многосточной (multi-sink). Стоки располагаются равномерно, т.е. количество конечных

сенсоров между двумя любыми соседними стоками одинаково. Первый и последний узел печочки являются стоками.

Разработана имитационная модель данной печочки для проведения симуляций на суперкомпьютере. За ее основу была взята модель Чиассерини–Гаретто. Эта модель была впервые предложена в [5] и является одной из первых, учитывающих наличие пассивной фазы (фазы сна) в режиме работы конечных устройств БСС. Наличие такой фазы, позволяющей устройствам более экономно расходовать заряд батарей, предусматривается, например, протоколом ZigBee.

Отметим, что применение суперкомпьютеров в моделировании весьма актуально на данный момент в связи со стремительным ростом масштабов сенсорных сетей (так, например, в строящемся «умном городе» PlanIT Valley в Португалии планируется довести количество сенсоров до 100 млн. [6]).

2. Основная часть

В нашей имитационной модели предполагается, что узлы БСС и среда передачи данных подчиняется предположениям модели Чиассерини–Гаретто [5]. Кратко перечислим основные положения этой модели. БСС представляется как совокупность некоторого числа одинаковых конечных сенсоров и одного стока. Радиус действия всех узлов сети предполагается одинаковым. Конечные устройства находятся в одном из двух состояний — активном либо пассивном. В активном состоянии осуществляется сбор, прием и передача информации, пассивный режим является энергосберегающим. Сток имеет постоянное питание, он всегда активен, но работает только на прием.

Время в модели дискретно. Длительности активной и пассивной фазы каждого конечного сенсора распределены по геометрическому закону с параметрами p и q соответственно. На каждом временном шаге конечные устройства генерируют несколько единиц информации, их количество распределено по закону Пуассона. Бuffer каждого устройства считается бесконечным.

После генерации данных сенсор стремится передать их соседнему устройству, который находится ближе к стоку. При этом возникает конкуренция за радиоканал, вызываемая интерференцией радиосигнала. Модель Чиассерини–Гаретто описывает эту конкуренцию с помощью следующих тезисов: 1) сенсор i может передать информацию сенсору j , если $d_{i,j} \leq r$ (где r — радиус действия сенсора); 2) для каждого другого принимающего сенсора k и передающего сенсора l должны быть выполнены следующие условия: $d_{i,k} > r$, $d_{l,j} > r$. Кроме того, запрещается одновременная передача и приём данных.

В данной работе на основе модели Чиассерини–Гаретто была разработана имитационная модель печочки сенсоров. Благодаря подбору расстояния между соседними сенсорами влияние интерференции сводится к минимуму: если i -й сенсор передает $i + 1$ -му, то $i - 1$ -й и $i + 1$ -й сенсор передавать на данном временном шаге не могут. Кроме того вводится дополнительное предположение, помогающее формализовать очередность захвата радиоканала. А именно, преимущество имеет тот сенсор, который сгенерировал на данном временном шаге меньше данных (он начинает процедуру захвата канала раньше).

Кратце каждый шаг модели состоит из трех этапов: 1) генерация данных конечными устройствами; 2) определение устройств, которые на данном шаге захватят канал; 3) передача информации между узлами сети.

Модель была реализована в виде программы на языке C++. Параллелизация проводилась с расчетом на вычислительную систему с распределенной памятью и коммуникации на основе MPI. Тестирование программы на суперкомпьютере BlueGene/P суперкомпьютерного комплекса МГУ им. М. В. Ломоносова показало хорошую масштабируемость модели. Это можно видеть по графику зависимости времени расчета от количества вычислительных узлов (см. рис. 1)

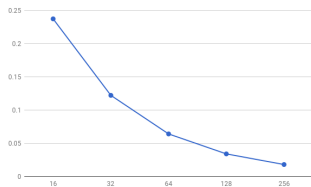


Рис. 1. Зависимость времени расчета от количества вычислительных узлов

3. Заключение

В рамках предлагаемой работы разработана имитационная модель, позволяющая исследовать с помощью суперкомпьютерной симуляции функционирование цепочки сенсоров с несколькими стоками. Тестирование на суперкомпьютере BlueGene/P показало, что модель обладает хорошей масштабируемостью. Описанная разработка может использоваться для моделирования работы систем мониторинга протяженных объектов, таких как автобаны или участки госграниц.

Литература

1. ZigBee Document 053474r06, Ver. 1.0, ZigBee Specification. ZigBee Alliance, 2004.
2. Schmid S., Wattenhofer R. Modeling sensor networks // Algorithms and Protocols for Wireless Sensor Networks / Ed. by A. Boukerche. — John Wiley & Sons, 2008.
3. Бузюков Л. Б., Окунева Д. В., Парамонов А. И. Проблемы построения беспроводных сенсорных сетей // Труды учебных заведений связи, 2017. — Т. 3, № 1.
4. Гащенко И. А. О моделировании беспроводных сенсорных сетей // International student research bulletin, Technical science. — 2017. — № 4.
5. Chiasserini C.-F., Garetto M. Modeling the Performance of Wireless Sensor Networks // Proc. Infocom, 2004.
6. URL: www.living-planit.com

UDC 004.72

Supercomputer Modelling of Wireless Sensor Network in Case of Connection Loss with Base Station

I. M. Nikolsky

*Faculty of Computational Mathematics and Cybernetics
Lomonosov Moscow State University
Leninskiye Gory, GSP-1, Moscow, 119991, Russian Federation*

Email: oliv_mail@mail.ru

The paper considers a chain of static sensors with multiple sinks. Nodes of the chain and the medium of data transmission are subject to assumptions of the Chiasserini-Garetto model. We take into account different features of wireless sensor networks (WSN) functioning like radio signal interference, presence of active and passive phases of end devices operation, etc. A parallel simulation model designed for evaluation of WSN parameters in absence of communication with the base station. The model is intended for computing systems with distributed memory and MPI communications. Testing on a BlueGene/P supercomputer shows good scalability of the model. The model enables investigation of various characteristics (e.g. speed of sink buffer filling) of WSNs used in the highway monitoring system.

Key words and phrases: wireless sensor network, imitational modelling, parallel algorithms, supercomputer.

УДК 004.4

Анализ производительности компьютерного моделирования физических процессов в системе длинных джозефсоновских переходов на кластере HybriLIT

А. Р. Рахронона*†

* Лаборатория информационных технологий
Объединённый институт ядерных исследований
ул. Жолио-Кюри, д. 6, Дубна, Московская область, Россия, 141980

† Кафедра системного анализа и управления
Государственный университет «Дубна»
ул. Университетская д. 19, Дубна, Московская область, Россия, 141982

Email: adiba-tj@inbox.rufirst@rudn.university

Исследована зависимость времени расчета вольт–амперной характеристики от число MPI процессов. Проанализировано влияние параметров вычисления на ускорение расчета. Показано, что с уменьшением шага по координате эффективность использования MPI технологии для расчета вольт–амперной характеристики увеличивается.

Ключевые слова: моделирование джозефсоновских переходов, MPI технология.

1. Введение

Слоистые высокотемпературные сверхпроводящие материалы могут быть рассмотрены как система связанных джозефсоновских переходов [1]. Интерес к математическому моделированию данной системы связан с богатыми нелинейными свойствами. Данная система демонстрирует ряд интересных физических явлений, таких как параметрический резонанс [2, 3], хаотические свойства [4], возникновение флоксонов [5, 6] и коллективных возбуждений [7]. Обнаруженное когерентное излучение из системы связанных джозефсоновских переходов дает возможность к различным практическим применениям [8, 9]. В частности, эта система является перспективным объектом сверхпроводящей электроники [10]. Все эти обстоятельства делают актуальным исследование физических явлений в системе связанных переходов. Поскольку динамика данной системы описывается нелинейными уравнениями, для их решения используют различные численные методы. Однако расчеты физических характеристик могут занимать достаточно большое время, что делает актуальным использование технологии параллельных вычислений.

В настоящей работе исследовано влияние параметров вычислительной схемы на время расчета вольт–амперной характеристики системы связанных джозефсоновских переходов. Вычисления проводились с использованием технологии параллельного программирования MPI на базе гетерогенного кластера HybriLIT.

2. Теоретическая модель и метод расчета

Для вычисления вольт–амперной характеристики системы связанных джозефсоновских переходов используется модель, которая учитывает индуктивную и емкостную связь между джозефсоновскими переходами [11, 12]. В рамках этой модели система дифференциальных уравнений для описания фазовой динамики связанных джозефсоновских переходов в нормированных единицах имеет следующий вид [11–13]

$$\begin{cases} \frac{\partial \varphi_l}{\partial t} = D_c V_l + s_c V_{l+1} + s_c V_{l-1}, \\ \frac{\partial V_l}{\partial t} = \sum_{k=1}^N \mathcal{L}_{lk}^{-1} \frac{\partial^2 \varphi_k}{\partial x^2} - \sin \varphi_l + \beta \frac{\partial \varphi_l}{\partial t} + I, \end{cases}$$

где φ_l и V_l разность фаз и напряжение в l -м ДП, соответственно. I — базовый ток нормированный на критический ток j_c , β — параметр диссипации, $D_c = 1 + (2\lambda_e/d_I) \coth(d_s/\lambda_e)$ — эффективная электрическая толщина ДП, нормированная на толщину диэлектрического слоя d_I , $s_c = -\lambda_e/[d_I \sinh(d_s/\lambda_e)]$ — параметр емкостной связи, d_s — толщина сверхпроводящего слоя, и λ_e — дебаевская длина проникновения. Напряжение нормировано на $V_0 = \hbar\omega_p/(2e)$, где $\omega_p = \sqrt{8\pi d_I j_c}/(\hbar e)$ — плазменная частота ДП, ε диэлектрическая проницаемость. Время t и координата x нормированы на значения плазменной частоты ω_p^{-1} и джозефсоновской глубины проникновения λ_J , соответственно. Матрица индуктивной связи $\hat{\mathcal{L}}$ имеет следующий вид

$$\hat{\mathcal{L}} = \begin{pmatrix} 1 & S & 0 & \cdots & & & S \\ & & & & \cdots & & \\ \cdots & 0 & S & 1 & S & 0 & \cdots \\ & & & & & & \\ S & & & & 0 & S & 1 \end{pmatrix},$$

где $S = s_\mathcal{L}/D_\mathcal{L}$ параметр индуктивной связи, $s_\mathcal{L} = -\lambda_L/\sinh(d_s/\lambda_L)$, $D_\mathcal{L} = d_I + 2\lambda_L \coth(d_s/\lambda_L)$ — эффективная магнитная толщина и λ_L лондоновская глубина проникновения. Отметим, что в рамках использованной нормировки параметр индуктивной связи S может принимать значения в интервале $S \in (-0.5, 0)$.

Алгоритм расчета ВАХ состоит из нескольких этапов. Вначале задается значение базового тока $I = 0$, также задается начальное и граничные условия. При $I = 0$ начальные условия для системы уравнений (2) имеют вид $\varphi_l(x, 0) = 0$ и $V_l(x, 0) = 0$. Граничные условия вдоль координаты x определяются значением внешнего магнитного поля $\partial\varphi_l/\partial x|_{x=0,L} = B_{ext}$, которое в нашем случае равно нулю. В направлении, перпендикулярном к плоскости слоев, использовано периодическое условие, т.е. при $l = N$, $\varphi_{l+1} = \varphi_1$, $V_{l+1} = V_1$, а при $l = 1$, $\varphi_{l-1} = \varphi_N$, $V_{l-1} = V_N$.

При фиксированном значении тока система уравнений (2) решается численно методом Рунге-Кутты четвертого порядка в интервале времени $[0, T_{max}]$ с шагом Δt . При этом производная по координате дискретизирована с шагом $\Delta x = 5\Delta t$ в интервале $[0, L]$. Как результат, решения системы уравнений определяются разность фаз φ_l и V_l как функция координаты и времени. Далее, необходимо вычислить среднее значение напряжения для заданного тока. При вычислении среднего значения напряжения интегрирования по координате проводилась по методу Симпсона и по времени методом прямоугольников. Таким образом, мы вычисляем соответствующее значение напряжения для заданного значения тока. Далее ток увеличивается на ΔI и вышеописанная процедура вычисления повторяется. Отметим, что полученные значения $\varphi_l(x, T_{max})$ и $V_l(x, T_{max})$ при I используются в качестве начальных условий при $I + \Delta I$. Более подробное описание процесса вычисления ВАХ приведено в работах [3, 13–15].

3. Численные результаты

Представленные здесь расчеты проводились с использованием параллельной программы, составленной на C++ с использованием технологии MPI. Параллельный алгоритм основан на блочном распределении расчетов в узлах по координате x на каждом шаге по времени. Подробное описание алгоритма приведено в [15]. Расчеты проводились на кластере HUBRIЛIT Многофункционального информационно-вычислительного комплекса Лаборатории информационных технологий ОИЯИ с использованием Intel-компилятора версии 17.0 (XE 2017).

На рис.1(а) представлена зависимость времени, затраченного для расчета ВАХ, от числа MPI-процессов для случая $\Delta x = 0.1$. Естественно, с увеличением MPI-процессов время расчета резко сокращается и, начиная с определенного количества

процессов, время расчета стабилизируется и далее почти не меняется с увеличением числа процессов. При использовании одного процесса для $\Delta x = 0.1$ время расчета составляло 1509 с, в то время при использовании 20 процессов оно составляло всего 258 с. Таким образом, для выбранного шага $\Delta x = 0.1$ мы получили ускорение расчета ВАХ примерно в 5.8 раза. Рис. 1(b) показывает зависимость времени расчета ВАХ для $\Delta x = 0.05$. Характер данной зависимости аналогичен показанному на Рис. 1(a). В связи с уменьшением шага по координате увеличивается время расчета по сравнению со случаем $\Delta x = 0.1$. В случае $\Delta x = 0.05$ вычисления с одним процессом длятся примерно 5280 с, а при использовании 20 процессов расчет занимает 652 с. Как видим, в этом случае мы получили 8-кратное ускорение вычислений при расчете в параллельном режиме.

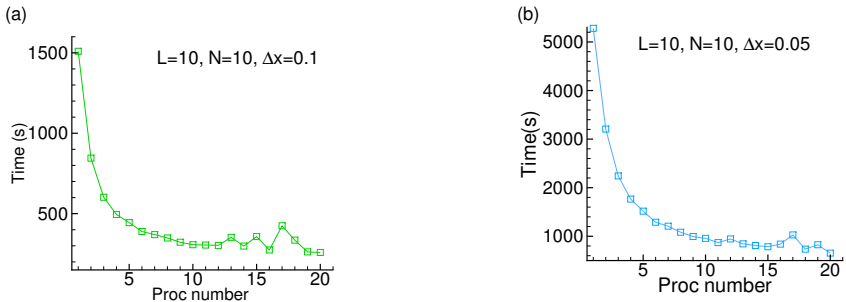


Рис. 1. (a) Зависимость времени вычисления от числа процессоров при шаге по координате $\Delta x = 0.1$; (b) То же, что в (a) при $\Delta x = 0.05$

Проведенный анализ показывает, что эффективность использования параллельных вычислений становится существенной при выборе маленьких шагов по координате.

4. Заключение

В заключение отметим, что нами исследовано влияние параметра шага дискретной сетки по координате на ускорение расчета ВАХ системы связанных джозефсоновских переходов при использовании MPI технологии. Представлена зависимость времени расчета ВАХ от числа MPI процессоров. Показано, что при уменьшении шага по координате эффективность использования MPI увеличивается.

Литература

1. Kleiner R., Steinmeyer F., Kunkel G., and Müller P. // Phys. Rev. Lett. — 1992. — Vol. 68. — P. 2394.
2. Shukrinov Yu. M. and Mahfouzi F. // Phys. Rev. Lett. — 2007. — Vol. 98. — P. 157001.
3. Rahmonov I. R., Shukrinov Y. M., and Irie A. // JETP Letters. — 2014. — Vol. 99. — P. 632.
4. Shukrinov Yu. M., Botha A. E., Medvedeva S. Yu., Kolahchi M. R., and Irie A. // Chaos. — 2014. — Vol. 24. — P. 033115.
5. Fulton T. A. and Dynes R. C. // Solid St. Commun. — 1972. — P. 57.
6. McLaughlin D. W. and Scott A. C. // Phys. Rev. A. — 1978. — Vol. 18. — P. 1652.

7. Kleiner R., Gaber T., and Hechtfischer G. // *Phys. Rev. B.* — 2000. — Vol. 62. — P. 4086.
8. Ozyuzer L., Koshelev A.E., Kurter C. et al. // *Science.* — 2007. — Vol. 318. — P. 1291.
9. Welp U., Kadowaki K. and Kleiner R. // *Nature Photonics.* — 2013. — Vol. 7. — P. 702.
10. Yurgens A. A. // *Supercond. Sci. Technol.* — 2000. — Vol. 13, R85.
11. Sakai S., Bodin P., and Pedersen N.F. // *J. Appl. Phys.* — 1993. — Vol. 73. — P. 2411.
12. Machida M. and Sakai S. // *Phys. Rev. B.* — 2004. — Vol. 70. — P. 144520.
13. Rahmonov I. R., Shukrinov Yu.M., Atanasova P.Kh., Zemlyanaya E.V., and Bashashin M.V. // *ЖЭТФ.* — 2017. — Vol. 124. — P. 131.
14. Shukrinov Yu., Rahmonov I., and Hamdipour M. // *LNCSC.* — 2012. — Vol. 7125. — P. 234.
15. Башашин М.В., Земляная Е.В., Рахмонов И.Р., Шукринов Ю.М., Атанасова П.Х., Волохова А.В. // *Компьютерные исследования и моделирование.* — 2016. — Vol. 8. — Pp. 593–604.

UDC 004.4

Analysis of the Performance of Computer Modeling of the Physical Processes in the System of Long Josephson Junctions in HybriLIT Cluster

A. R. Rahmonova*[†]

** Laboratory of Information Technologies
Joint Institute for Nuclear Research*

6 Joliot-Curie, Dubna, Moscow Region, 141980, Russian Federation

*† Department of System Analysis and Management
Dubna State University*

19, Universitetskaya str., Dubna, Moscow Region, 141982, Russian Federation

Email: adiba-tj@inbox.ru first@rudn.university

The dependence of the calculation time of current–voltage characteristic on the number of MPI processes is investigated. The influence of the parameters of simulation on the acceleration of calculation is analysed. It is demonstrated that, with the decreasing of the coordinate step size the efficiency of using of MPI technology for the calculation of the current–voltage characteristic is increased.

Key words and phrases: modeling of Josephson junction, MPI technology.

УДК 004.4

Разработка сервиса $VM@N$ -Webapp для обработки и визуализации информации об эксперименте на фиксированной мишени $VM@N$

А. И. Чеботов^{‡§}, Р. Л. Авзалов^{‡§}, Я. М. Чернышев^{‡§},
К. В. Герценбергер[†], Ю. А. Бутенко*

* Лаборатория информационных технологий
Объединённый институт ядерных исследований
ул. Жолио-Кюри, д. 6, Дубна, Московская область, Россия, 141980

† Лаборатория физики высоких энергий
Объединённый институт ядерных исследований
ул. Жолио-Кюри, д. 6, Дубна, Московская область, Россия, 141980

‡ Учебно научный центр
Объединённый институт ядерных исследований
ул. Жолио-Кюри, д. 6, Дубна, Московская область, Россия, 141980

§ Государственный университет «Дубна»
ул. Университетская, д. 19, Дубна, Московская область, 141982

Email: alchebotov@gmail.com, rustam.avzalov@yandex.ru, chyam.995@gmail.com, gertsen@jinr.ru, ybutenko@jinr.ru

Для удобного просмотра, изменения и визуализации информации о проводимом в Объединённом институте ядерных исследований (ОИЯИ) эксперименте $VM@N$ появилась необходимость разработки веб-сервиса $VM@N$ -WebApp. Реализуемый сервис предназначен для работы с унифицированной базой данных эксперимента, заполненной параметрами проведённых сеансов эксперимента $VM@N$, которые в дальнейшем активно используются членами коллаборации для обработки экспериментальных (а также смоделированных) данных, в том числе для реконструкции и анализа получаемых данных. В статье приведён обзор данного сервиса.

Ключевые слова: веб-приложение, сервис, база данных, физический эксперимент.

1. Введение

Исследование свойств ядерной материи в экстремальных условиях - одно из наиболее интенсивно развивающихся направлений современной физики. Эксперименты по столкновению тяжелых ионов, в которых создаются условия сверхвысоких температур и плотностей, интенсивно ведутся или планируются в различных научных центрах Европы и США. Для проведения этих исследований сейчас создается ускорительно-экспериментальный комплекс $NICA$ (ОИЯИ, Дубна, Россия) с целью изучения свойств плотной барионной материи. На первом этапе проекта $NICA$ реализуется эксперимент на фиксированной мишени $VM@N$ [1]. В ходе экспериментов на физической установке $VM@N$ будут изучаться взаимодействия ядер пучка с ядрами мишени. Сервис $VM@N$ -WebApp представляет собой веб-приложение, предназначенное как для пользователей системы - членов коллаборации эксперимента $VM@N$, предоставляя информацию о полученных экспериментальных данных, так и для группы администрирования сервиса.

Сервис состоит из двух основных модулей (см. рис. 1): $VM@N$ WebApp и API. $VM@N$ WebApp [2] - веб-приложение, с которым непосредственно работают пользователи сервиса. Оно содержит все веб-ресурсы сервиса, из которых формирует затем веб-страницы, ограничивая доступ к контенту в зависимости от сертификатов, полученных от другого модуля - API.

API [3] помимо сертификатов доступа к контенту передает на $VM@N$ WebApp всю информацию, хранимую в базе данных (данные запуска экспериментов). API имеет интерфейсы взаимодействия с единой системой управления идентификацией пользователей сервиса $VM@N$ WebApp - FreeIPA, и базой данных, реализованной при

1. Запись экспериментальных данных эксперимента *BM@N*. При проведении сеансов эксперимента *BM@N* производится запись получаемых данных в бинарные файлы разработанного формата *mpd*.
2. Сохранение информации об эксперименте *BM@N* в базу данных *PostgreSQL*. Помимо информации об эксперименте, такой как: время проведения сеансов, энергии столкновения частиц с мишенью, получаемых файлах с данными и многое другое, – также в базу данных эксперимента сохраняются файлы с геометрией детекторов установки, которые представляют собой *ROOT*-файлы. *API* получает хранимые данные из базы данных, используя *Node.js*. Файл из базы данных представляет, собой массив двоичных данных, сохранённый в *ROOT* файл в распределённой файловой системе *EOS* [9].
3. Визуализация геометрии установки с помощью библиотеки *jsROOT*. Чтобы приступить к этапу визуализации необходимо данные из *ROOT* файла конвертировать в *JSON* - формат с помощью метода *TBufferJSON::ConvertToJSON*. Далее *JSON* файл отправляется на *API* по средствам *Node.js*, далее на веб-приложение, написанное на *Angular*.

4. Дальнейшее развитие

В дальнейшем планируется развитие проекта по многим аспектам: безопасность, стабильность и добавление нового функционала. Ближайшие задачи:

- Создание панели управления доступом для администратора с помощью *LDAPjs*.
- Разработка системы визуализации хранимых данных, представленных *ROOT*-деревьями, с помощью библиотеки *jsROOT*.

5. Заключение

В статье представлен обзор разрабатываемого веб-сервиса *BM@N WebApp*. Сервис ориентирован на предоставление членам коллаборации эксперимента *BM@N* информации о проведённых сеансах, полученных экспериментальных данных и другие параметры, используемые в дальнейшем при обработке и анализе. Веб-сервис в настоящее время активно развивается, дополняется новыми возможностями и функциями.

Литература

1. Проект *BM@N*. <http://jinrmg.jinr.ru/koi8/2012/19/bm19.html>.
2. *BM@N WebApp*. <https://gitlab-hlit.jinr.ru/nica/bman-webapp.git>.
3. *API*. <https://gitlab-hlit.jinr.ru/hybrilit/hlit-api>.
4. *NodeJS*. <https://nodejs.org/en/>.
5. *Angular*. <https://angular.io/>.
6. *ROOT*. <https://root.cern.ch/>.
7. *JavaScript ROOT*. <https://github.com/root-project/jsroot/blob/master/docs/JSROOT.md>.
8. *D3.js*. <https://d3js.org/>.
9. *EOS-Storage*. <https://eos.readthedocs.io/en/latest/>.

UDC 004.4

Development of BM@N-Webapp Service for Processing and Visualization of Information about the Experiment on a Fixed Target BM@N

A. I. Chebotov^{†§}, R. L. Avzalov^{†§}, Y. M. Chernyshev^{†§},
K. V. Gertsenberger[†], Y. A. Butenko*

* *Laboratory of Information Technologies*

Joint Institute for Nuclear Research

6 Joliot-Curie, Dubna, Moscow region, 141980, Russian Federation

† *Laboratory of High Energy Physics*

Joint Institute for Nuclear Research

6 Joliot-Curie, Dubna, Moscow region, 141980, Russian Federation

‡ *University Centre*

Joint Institute for Nuclear Research

6 Joliot-Curie, Dubna, Moscow region, 141980, Russian Federation

§ *State University "Dubna"*

19 University, Dubna, Moscow region, 141982, Russian Federation

Email: alchebotov@gmail.com, rustam.avzalov@yandex.ru, chyam.995@gmail.com, gertsen@jinr.ru, ybutenko@jinr.ru

For the convenient viewing, changing and visualization of information about the *BM@N* experiment conducted at the Joint Institute for Nuclear Research (*JINR*), the necessity of developing a web service *BM@N-WebApp* has appeared. The implemented service is designed to work with a unified experimental database filled with the parameters of the conducted *BM@N* experiment sessions, which are further actively used by the collaboration members to process the experimental (as well as modeled) data, including for reconstruction and analysis of the received data. The article gives an overview of this service.

Key words and phrases: web application, service, database, physical experiment.

Математическое моделирование
Mathematical modeling

UDC 519.864.3

Optimal Multi-Period Self-Financing Portfolio Strategy with Commission

M. S. Al-Nator, S. V. Al-Nator

*Department of Data Analysis, Decision Making, and Financial Technology
Financial University under the Government of Russian Federation
40 Leningradsky Prospekt, Moscow, 125993, GSP-3, Russian Federation*

Email: malnator@yandex.ru, salnator@yandex.ru

The problem of choosing the optimal multi-period self-financing portfolio strategy with discrete time and commission is solved. It is assumed that all the released money from the sale of some assets will be invested in the acquisition of other assets in order to maximize the portfolio value at the end of the investment horizon.

Key words and phrases: portfolio, long position, multi-period portfolio transaction, optimal strategy, commission.

1. Introduction

The goal of any financial transaction in the securities market is to achieve maximum income and to increase the initial capital. In this work for the Markowitz model (only long positions are allowed for this model, see [1, 2]) we solve the problem of choosing the optimal multi-period self-financing portfolio strategy with commission under certainty. In other words, we are looking for a strategy for which all the released money from the sale of some assets will be invested in the acquisition of other assets in order to maximize the portfolio value at the end of the investment horizon.

We emphasize that the solution of this problem is nontrivial and may differ from its solution without commission.

Under uncertainty the financial analyst usually considers the most likely scenarios of the possible completion of the transaction. In that case, the analysis of individual scenarios is carried out under certainty. The condition of certainty means that the investor knows both the current and future prices (based on price forecasts) of assets and income. The investigation of portfolio transactions under certainty is definitely useful and widely used by accountants and auditors in the analysis of closed transactions.

One-period portfolio analysis with commission under uncertainty was investigated in detail in [3–5].

2. Preliminaries and notation

Let the market consist of n assets A_1, A_2, \dots, A_n . The portfolio at time t will be denoted by the position vector $\mathbf{z}(t) = (z_1(t), z_2(t), \dots, z_n(t))$, where $z_k(t) \geq 0$ is the position of A_k (note that $z_k(t)$ is the amount of A_k in the portfolio at time t). Let $p_k(t)$ be the asset price of A_k at time t . Then the market state at any time t is specified by the n -dimensional price vector $\mathbf{p}(t) = (p_1(t), p_2(t), \dots, p_n(t))$.

In what follows, we consider multi-period portfolio transactions with fixed and finite investment horizon. We assume that changes in asset prices occur only at discrete instants of time.

By strategy we mean a sequential restructuring of the portfolio (formed at the time $t_0 = 0$ with an initial capital) at the moments $t = t_0, t_1, \dots, t_N$, in order to maximize the portfolio value at the time t_N . We denote the strategy by

$$\mathbf{Z}_{[t_0, t_N]} = \{\mathbf{z}(t_0), \mathbf{z}(t_1), \dots, \mathbf{z}(t_{N-1})\}.$$

The strategy $\mathbf{Z}_{[t_0, t_N]}$ is called admissible if all the released money from the sale of some assets is invested in the acquisition of other assets.

3. Main results

The optimization problem with commission α is formulated as follows:

For the investment horizon $[t_0, t_N]$ and the price forecast $[p_k(t_i)]_{k=1, n}^{i=0, N}$ find a strategy

$$\mathbf{Z}_{[t_0, t_N]}^* = \{\mathbf{z}^*(t_0), \mathbf{z}^*(t_1), \dots, \mathbf{z}^*(t_{N-1})\}$$

that satisfies at any time $t_i, i = 1, \dots, N$ the balance equation

$$\sum_{k=1}^n z_k(t_{i-1})p_k(t_i) = \sum_{k=1}^n z_k(t_i)p_k(t_i) + \alpha \sum_{k=1}^n |z_k(t_i) - z_k(t_{i-1})|p_k(t_i)$$

such that

$$\langle \mathbf{z}^*(t_{N-1}), \mathbf{p}(t_N) \rangle = \max_{\mathbf{Z}_{[t_0, t_N]}} \langle \mathbf{z}(t_{N-1}), \mathbf{p}(t_N) \rangle,$$

where $\langle \cdot, \cdot \rangle$ is the standard scalar product in \mathbb{R}^n .

We need the following notation to describe the optimal strategy construction scheme:

$$c_k(t_i, t_{i+1}) = \frac{p_k(t_{i+1})}{p_k(t_i)}, \quad c_{\max}(t_i, t_{i+1}) = \max_{1 \leq k \leq n} c_k(t_i, t_{i+1}), \quad k = \overline{1, n}, \quad i = \overline{0, N-1},$$

$$S(t_i) = \sum_{k=1}^n z_k(t_i)p_k(t_i), \quad i = \overline{0, N},$$

$$\lambda_k(t_i) = \frac{S_k(t_N)}{S_k(t_i)}, \quad k = \overline{1, n}, \quad i = \overline{0, N},$$

$$\omega_k(t_i, t_{i+1}) = \lambda_k(t_{i+1})c_k(t_i, t_{i+1}), \quad k = \overline{1, n}, \quad i = \overline{0, N-1},$$

$$\omega_{\max}(t_i, t_{i+1}) = \max_{1 \leq k \leq n} \omega_k(t_i, t_{i+1}), \quad i = \overline{0, N-1},$$

$$\rho_k(t_i, t_{i+1}) = a_\alpha \omega_{\max}(t_i, t_{i+1}) - \omega_k(t_i, t_{i+1}), \quad k = \overline{1, n}, \quad i = \overline{0, N-1},$$

where

$$a_\alpha = \frac{1 - \alpha}{1 + \alpha}.$$

The optimal strategy is constructed in the direction from the end of the investment horizon to its beginning:

1. For each t_i , select the asset A_m with $\omega_m(t_i, t_{i+1}) = \omega_{\max}(t_i, t_{i+1})$.

2. For each $k = \overline{1, n}$ calculate $\rho_k(t_i, t_{i+1})$.

2.1. If $\rho_k(t_i, t_{i+1}) \leq 0$ then the position $z_k(t_{i-1})$ of A_k is not changed under the transition $\mathbf{z}(t_{i-1}) \rightarrow \mathbf{z}(t_i)$. In that case set $\lambda_k(t_i) = \omega_k(t_i, t_{i+1}) = c_k(t_i, t_{i+1})\lambda_k(t_{i+1})$.

2.2. If $\rho_k(t_i, t_{i+1}) > 0$ then sell all units of the asset A_k and invest the money from its sale in the asset A_m . In that case set $\lambda_k(t_i) = a_\alpha \omega_{\max}(t_i, t_{i+1})$.

Acknowledgments

The authors are grateful to Professor Kasimov Yu. F. for useful discussions.

References

1. M. S. Al-Nator, Yu. F. Kasimov, A. N. Kolesnikov, Fundamentals of financial calculations, part 2 [in Russian], Financial University Publishing House, Moscow, 2013.
2. Yu. F. Kasimov, M. S. Al-Nator, A. N. Kolesnikov, Fundamentals of financial calculations. Basic financial transaction calculation schemes [in Russian], Knorus, Moscow, 2017.
3. M. S. Al-Nator, S. V. Al-Nator, Yu. F. Kasimov: Portfolio analysis with transaction costs Under Uncertainty, J. Math. Sci. (2016) 214 (1), 12–21. doi:110.1007/s10958-016-2754-9.
4. M. S. Al-Nator: On the choice of an optimal portfolio with commission for the Markowitz model, in Materials of VI All-Russian conference with international participation Information and telecommunication technologies and mathematical simulation of hi-tech systems [in Russian], 2016, Moscow, RUDN University. P. 214–216.
5. M. S. Al-Nator: Portfolio analysis with general commission, in Materials of XXXIII International Seminar on Stability Problems for Stochastic Models, OP&PM, 2016, Vol. 23, Issue 2. P. 140–141.

УДК 519.864.3

Оптимальная самофинансируемая многопериодная портфельная стратегия с учетом комиссии

М. С. Аль-Натор, С. В. Аль-Натор

*Департамент анализа данных, принятия решений и финансовых технологий
Финансовый университет при Правительстве РФ
Ленинградский просп., д. 49, ГСП-3, Москва, Россия, 125993*

Email: malnator@yandex.ru, salnator@yandex.ru

Решена задача о выборе оптимальной стратегии в многопериодной самофинансируемой портфельной сделке с дискретным временем и с учетом комиссии. Предполагается, что все освободившиеся средства от продажи некоторых активов вкладываются в другие активы с целью максимизации стоимости портфеля в конце инвестиционного горизонта.

Ключевые слова: портфель, длинная позиция, оптимальная стратегия, многопериодная портфельная сделка, оптимальная стратегия, комиссия.

УДК 519.633, 538.9

Об одном методе приближенного решения задачи трех тел**И. В. Амирханов, Н. Р. Саркар**

*Лаборатория информационных технологий
Объединенный институт ядерных исследований
ул. Жолио-Кюри, д. 6, Дубна, Московская область, Россия, 141980*

Email: camir@jinr.ru, sarker@jinr.ru

В работе предложен метод исследования решений одной квантово-механической задачи трех тел. Исследована краевая задача для системы из двух частиц, движущихся в потенциальном поле третьей частицы. Потенциалы взаимодействия между частицами выбраны квадратично растущими. В этом случае задача двух тел имеет точное решение. Решение исходной задачи трех тел представлено в виде разложения по решениям задачи двух тел. На коэффициенты разложения получается однородное линейное матричное уравнение. Приравняв определитель этого уравнения к нулю, находим собственные значения исходной задачи и коэффициенты разложения. Так как элементы матрицы алгебраических уравнений с решениями задачи двух тел вычисляются аналитически, то собственные значения задачи трех тел явным образом зависят от параметров потенциалов. Меняя эти параметры произвольным образом, можно получить желаемые спектры, т.е. можно управлять собственными значениями исходной задачи. Например, если выключить взаимодействие между первой и второй частицей, то получаем точные решения двух не взаимодействующих частиц в поле третьей частицы. Выше сказанные утверждения подтверждаются при рассмотрении в рядах частных случаев при решении исходной задачи.

Ключевые слова: задача трех тел, растущий потенциал, собственные значения.

1. Введение

При исследовании решений многочастичных задач часто бывает удобным сводя эту задачу к решениям задачи трех тел. Подобные задачи возникают при изучении энергетических состояний системы, состоящей из двух частиц, движущихся в потенциальном поле третьей частицы или самосогласованном поле других частиц [1, 2]. К таким системам относится [3] атом He , содержащий два электрона, движущихся в кулоновском поле ядра с зарядом $Z = 2$.

В данной работе предлагаем метод решения подобной задачи с квадратично растущим потенциалом. Пренебрегая спин-орбитальным взаимодействием можно записать оператор Гамильтона системы в виде

$$H = H_1 + H_2 + V_{12},$$

где $H_1 = -\frac{\hbar^2}{2\mu_1}\nabla_1^2 + C_1r_1^2$, $H_2 = -\frac{\hbar^2}{2\mu_2}\nabla_2^2 + C_2r_2^2$, $V_{12}(|\vec{r}_1 - \vec{r}_2|) = C_3|\vec{r}_1 - \vec{r}_2|^2$, H_1, H_2 — оператор Гамильтона двух частиц с квадратично растущим потенциалом, V_{12} — оператор взаимодействия между частицами, C_1, C_2 и C_3 — параметры задачи.

2. Постановка задачи

Приближенные решения задачи трех тел сводятся к нахождению собственных функций и собственных значений стационарного уравнения Шредингера

$$H\Psi = E\Psi, \quad (1)$$

при условии сохранения нормировки

$$\int d\vec{r}_1 d\vec{r}_2 \left| \psi(\vec{r}_1, \vec{r}_2) \right|^2 = 1. \quad (2)$$

Задача двух тел

$$H_1 \varphi_{1n}(\vec{r}_1) = E_{1n} \varphi_{1n}(\vec{r}_1), \quad H_2 \varphi_{2m}(\vec{r}_2) = E_{2m} \varphi_{2m}(\vec{r}_2), \quad n, m = 1, 2, 3, \dots \quad (3)$$

имеет точные решения. Так как решения φ_{1n} и φ_{2m} являются полным набором ортонормированных функций, то решения уравнения (3) будем искать в виде

$$\Psi = \sum_{n,m=1}^{N,M} B_{nm} \varphi_{1n} \varphi_{2m}, \quad (4)$$

где N, M — максимальные количества членов, удерживаемые в разложении.

В качестве $\varphi_{1n}(r_1)$ и $\varphi_{2m}(r_2)$ выбираем сферические симметричные решения, т.е. $l = 0$. Подставляя (4) в (1) получаем однородную систему алгебраических уравнений для нахождения E и B_{nm}

$$[E - E_{1n} - E_{2m}] B_{nm} = \sum_{n',m'=1}^{N,M} W_{nmm'n'} B_{n'm'},$$

где $W_{nmm'n'} = \int d\vec{r}_1 \int d\vec{r}_2 \varphi_{1n}(r_1) \varphi_{2m}(r_2) V_{12} \varphi_{1n'}(r_1) \varphi_{2m'}(r_2)$.

После интегрирования по углам получаем

$$[E - E_{1n} - E_{2m}] B_{nm} = C_3 \left[\sum_{m'} A_{mm'} B_{nm'} + \sum_{n'} D_{nn'} B_{n'm'} \right], \quad (5)$$

где $A_{mm'} = \int_0^\infty dr_2 r_2^2 \varphi_{2m}(r_2) \varphi_{2m'}(r_2)$, $D_{nn'} = \int_0^\infty dr_1 r_1^2 \varphi_{1n}(r_1) \varphi_{1n'}(r_1)$.

Для вычисления $A_{mm'}$ и $D_{nn'}$ приведем явный вид φ_{1n} и φ_{2m} . Несколько нижних собственных функций φ_{1n} и собственных значений E_{1n} ($n = 1, 2, 3$) имеет вид [4]:

$$\varphi_{11} = N_{11} r \exp^{-\beta_1 r^2},$$

$$\varphi_{12} = N_{12} r(r^2 - a_1) \exp^{-\beta_1 r^2}, \quad \varphi_{13} = N_{13} r(r^2 - a_2)(r^2 - a_3) \exp^{-\beta_1 r^2},$$

где N_{1i} ($i = 1, 2, 3$) — нормировочные постоянные,

$$E_{11} = 3\sqrt{C_1}, E_{12} = 7\sqrt{C_1}, E_{13} = 11\sqrt{C_1}, a_1 = 3/2\sqrt{C_1}, a_2 = (5 - \sqrt{10})/2\sqrt{C_1},$$

$$a_3 = (5 + \sqrt{10})/2\sqrt{C_1}, \beta_1 = \sqrt{C_1}/2, D_{11} = 3/2\sqrt{C_1}, D_{22} = 7/2\sqrt{C_1}, D_{33} = 11/2\sqrt{C_1},$$

$$D_{12} = D_{21} = \sqrt{3/2C_1}, D_{23} = D_{32} = \sqrt{5/2C_1}.$$

Решения φ_{2m} и $A_{mm'}$ получаются заменой C_1 на C_2 .

Рассмотрим ряд частных случаев

1. Пусть $M = N = 1$, тогда решение системы (5) будут иметь вид

$$E = 3\sqrt{C_1} + 3\sqrt{C_2} + C_3(3/(2\sqrt{C_1}) + 3/(2\sqrt{C_2})).$$

2. Пусть $N = 1$, $M = 2$, тогда два решения системы (5) будут иметь вид

$$E_{1,2} = p/2 \pm \sqrt{p^2/4 - q},$$

где

$$p = \gamma_1 + \gamma_2; \quad q = (\gamma_1\gamma_2 - \gamma_3), \quad \gamma_1 = 3\sqrt{C_1} + 3\sqrt{C_2} + C_3(3/(2\sqrt{C_1}) + 3/(2\sqrt{C_2})), \\ \gamma_2 = 3\sqrt{C_1} + 7\sqrt{C_2} + C_3(3/2\sqrt{C_1} + 7/2\sqrt{C_2}), \quad \gamma_3 = 3/2C_1.$$

3. Пусть $N = 2$, $M = 2$, тогда получаем четыре решения и т.д.
Коэффициенты B_{nm} находим с учетом условия нормировки (2).

3. Выводы

1. Собственные значения E явным образом выражаются через C_1 , C_2 и C_3 .
2. При $C_3 \rightarrow 0$ получаем точные решения двух невзаимодействующих частиц в поле третьей частицы.

Литература

1. Каширина Н. И., Лахно В. Д., Математическое моделирование автолокализованных состояний в конденсированных средах. М.: Физматлит, 2013. — 292 с.
2. Амирханов И. В., Пузынин И. В., Стриж Т. А., Лахно В. Д. Решение уравнений ЛПП в теории биполярона // Известия АН, серия физическая. 1995. — Т. 59, № 8. — С. 106–110.
3. Давидов А. С. Квантовая механика (2-е изд.). — М.: Наука, 1973.
4. Амирханов И. В., Саркар Н. Р. Об одном методе исследования самосогласованной нелинейной краевой задачи на собственные значения с растущими потенциалами // Вестник РУДН. Серия: Математика. Информатика. Физика. — 2018. — Т. 26, № 1. — С. 51–59.

UDC 519.633, 538.9

On a Method of Approximate Solution of the Three-Body Problem

I. V. Amirkhanov, N. R. Sarker

*Laboratory of Information Technologies
Joint Institute for Nuclear Research*

6 Joliot-Curie Str., Dubna, Moscow Region, Russian Federation, 141980

Email: camir@jinr.ru, sarker@jinr.ru

A method is proposed for investigating the solutions of a single quantum-mechanical three-body problem. The boundary value problem for a system of two particles moving in the potential field of a third particle is investigated. The interaction potentials between the particles are chosen to be quadratically growing. In this case the two-body problem has an exact solution. The solution of the original three-body problem is represented as an expansion in the solutions of the two-body problem. The coefficients of the expansion yield a homogeneous linear matrix equation. Equating the determinant of this equation to zero, we find the eigenvalues of the original problem and the coefficients of the expansion. Since the elements of the matrix of algebraic equations with solutions of the two-body problem are calculated analytically, the eigenvalues of the three-body problem depend explicitly on the parameters of the potentials. Changing these parameters in an arbitrary way, we can obtain the desired spectra, i.e. you can control the eigenvalues of the original problem. For example, if turning off the interaction between the first and second particles, then we obtain exact solutions of two noninteracting particles in the field of the third particle. The above statements are confirmed when they are considered in the series of particular cases in the solution of the original problem.

Key words and phrases: three-body problem, growing potential, eigenvalues.

УДК 004.021, 517.938, 519.2

Простая модель системы с активным управлением очередью по алгоритму типа RED

А.-М. Ю. Апреутесей

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклухо-Маклая, д.6, Москва, Россия, 117198*

Email: 1032152610@pfur.ru

В данной работе с помощью средства моделирования OpenModelica реализована упрощенная модель системы с активным управлением очередью с алгоритмом типа Random Early Detection (RED), а так же продемонстрирована динамика поведения основных параметров данной системы.

Ключевые слова: алгоритм случайного раннего обнаружения (RED), фазовый портрет, OpenModelica.

1. Введение

Алгоритм активного управления очередью (Active Queue Management, AQM) с алгоритмом управления типа Random Early Detection (RED) обеспечивает контроль и предотвращение перегрузок в очередях маршрутизаторов [1]. В модели такой системы используются дифференциальные включения, что создает сложности в процессе исследования. При имитационном моделировании оптимально использовать гибридный подход, который можно реализовать на языке Modelica. Modelica является объектноориентированным языком физического моделирования, подходящим для имитации как непрерывных, так и дискретных систем. Целью данной работы является моделирование упрощенной модели системы с управлением по алгоритму RED и демонстрация поведения основных ее параметров.

2. Упрощенная модель системы с управление по алгоритму типа RED

Модель AQM с управлением по алгоритму типа RED представима в виде трех уравнений:

$$\dot{W}(t) = \frac{1}{R} - \frac{W(t)W(t-R)}{2R(t-R)} KQ(t-R), \quad (1)$$

$$\dot{Q}(t) = \begin{cases} N \frac{W(t)}{R} - C, & Q(t) > 0, \\ \max \left\{ N \frac{W(t)}{R} - C, 0 \right\}, & Q(t) = 0, \end{cases} \quad (2)$$

$$\dot{q}(t) = \frac{\ln(1-\alpha)}{\delta} \hat{q}(t) - \frac{\ln(1-\alpha)}{\delta} Q(t), \quad (3)$$

описывающих поведение трех функций: $W(t)$ — средний размер TCP окна (в пакетах); $Q(t)$ — средний размер очереди (в пакетах); $\hat{q}(t)$ — экспоненциально взвешенное скользящее среднее значение длины очереди.

Рассматриваемая система содержит C приборов, что соответствует скорости обработки пакетов в очереди (C пакетов в единицу времени). Так же на средний размер TCP окна влияет вероятностная функция сброса пакета $p(\cdot)$ (собственно управление). В системе так же учитываются такие параметры, как $N(t)$ — число TCP-сессий и $R(t)$ — время двойного оборота (в сек) (Round Trip Time, RTT). Полная модель представлена в статье [2].

Положим, что функция сброса пакетов пропорциональна длине очереди $Q(t)$, т.е. зададим $p(\cdot) = KQ(t)$, где K — коэффициент пропорциональности. Более того, примем, что такие параметры как $N(t)$ и $R(t)$ являются постоянными, не изменяющимися со временем, т.е. $N(t) \equiv N$ и $R(t) \equiv R$.

Уравнение (1) описывает динамическое изменение среднего размера окна TCP. Левое слагаемое отражает фазу медленного старта, во время которой размер окна увеличивается за единицу времени двойного оборота, правое же слагаемое описывает мультипликативное уменьшение размера окна в два раза при каждой потере пакета.

Рассмотрим уравнение (2), описывающее поведение среднего размера очереди. Компонента $NW(t)/R$ отражает среднюю интенсивность поступления пакетов, что соответствует увеличению длины очереди при поступлении пакетов. Разность средней интенсивности поступления пакетов и пропускной способности C моделирует уменьшение длины очереди в случае, если она больше нуля.

Уравнение (3) описывает экспоненциально взвешенное скользящее среднее значение длины очереди, где δ — время между поступлениями пакетов в систему: $\delta = 1/C$, α — весовой коэффициент экспоненциально взвешенного скользящего среднего размера длины очереди: $\alpha = 1 - e^{-1/C}$.

В системе OpenModelica программа имеет вид:

```
model TCP_RED
extends Modelica.Math;
parameter Real N(start=1), K(start=5.3), C(start=1), R(start=1);
Real alfa, q_cap(start=0.1), q(start=1), Real w(start=0.1);
equation
  alfa=1-exp(-1/C);
  der(q_cap)=(log(1-alfa)*q_cap/(1/C))-(log(1-alfa)*q/(1/C));
  der(w)=(1/R)-(w*delay(w,R)*K*delay(q,R)/(2*R));
  der(q)= if q>0 then ((N*w/R)-C) else max(((N*w)/R)-C,0);
end TCP_RED;
```

При моделировании данной системы с начальными значениями параметров $N = 1$, $R = 1$, $K = 5, 3$, $W(0) = 0, 1$, $Q(0) = 1$, $\hat{q}(t) = 0, 1$ были получены графики динамики изменения $W(t)$, $Q(t)$, $\hat{q}(t)$ (рис. 1) и фазовый портрет (W, Q) . Фазовая траектория осциллирует вокруг своей стационарной точки (рис. 2), что отражает наличие автоколебаний параметров системы [3].

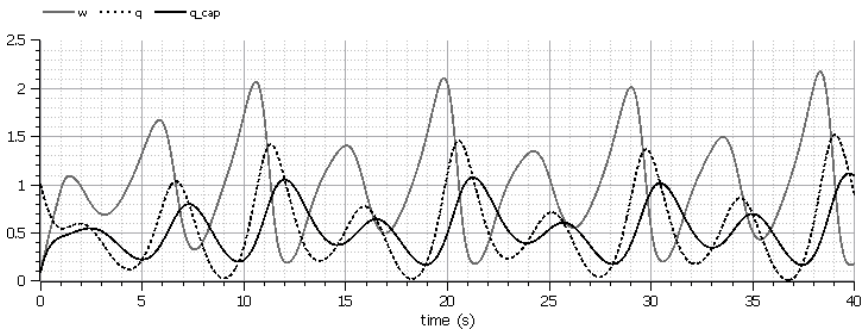


Рис. 1. Динамика изменения $W(t)$, $Q(t)$ и $\hat{q}(t)$

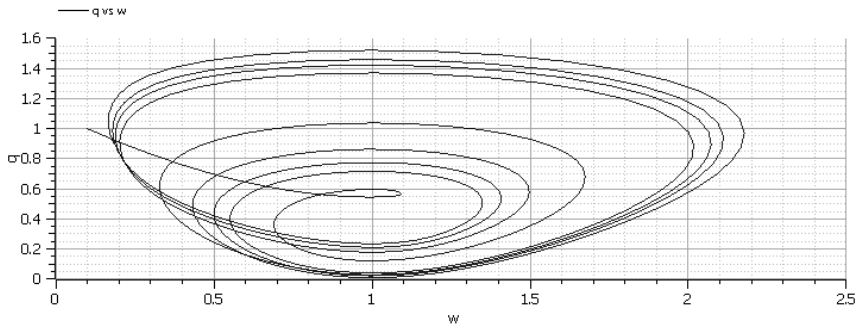


Рис. 2. Фазовый портрет (W, Q)

3. Заключение

В работе продемонстрировано применение языка Modelica к моделированию гибридной простой системы с управлением по алгоритму типа RED. Показано наличие автоколебательного режима при определенных заданных параметрах системы.

Литература

1. Floyd S., Jacobson V. Random Early Detection Gateways for Congestion Avoidance // IEEE/ACM Transactions on Networking. — 1993. — Vol. 1, Issue 4. — P. 397–413.
2. Misra V., Gong W.-B., Towsley D. Fluid-Based Analysis of a Network of AQM Routers Supporting TCP Flows with an Application to RED // ACM SIGCOMM Computer Communication Review. — Volume 30, Issue 4, October 2000. — P. 151–160.
3. Королькова А. В. Определение области возникновения автоколебаний в системах типа RED // Вестник РУДН. Серия «Математика. Информатика. Физика». — 2010. — № 1. — С. 103–105.

UDC 004.021, 517.938, 519.2

A Simple Model of Active Queue Management System According to the RED Algorithm

A.-M. Y. Apreutesey

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
Miklukho-Maklaya str. 6, Moscow, 117198, Russian Federation*

Email: 1032152610@pfur.ru

Using a modeling tool such as OpenModelica a simplified model of Active Queue Management system with Random Early Detection (RED) algorithm is constructed. The behaviour of main system parameters is demonstrated.

Key words and phrases: Random Early Detection (RED), phase portrait, OpenModelica.

УДК 519.218, 519.22

Вероятностные и статистические методы, применяемые для оптимизации работы парковочных систем и связанные с распознаванием номерных знаков

Д. М. Ахметханова*, И. С. Зарядов*†

* Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклухо-Маклая, д. 6, Москва, Россия, 117198

† Институт проблем информатики
Федеральный исследовательский центр «Информатика и управление» РАН
ул. Вавилова, д. 44, к. 2, Москва, Россия, 119333

Email: 1032143094@rudn.university, zaryadov_is@rudn.university

Формулируются основные вероятностные и статистические методы, которые могут применяться при решении задач оптимизации работы парковочной системы, а также связанные с распознаванием номерных знаков.

Ключевые слова: автоматизированная парковочная система, распознавание номерных знаков, регрессионные модели, скрытые марковские модели, оптимизационные задачи в теории массового обслуживания.

1. Введение

Количество автомобилей растет изо дня в день, а с ними и актуальность автоматизированных платных парковок. Автоматизированная парковочная система (кратко АПС) представляет собой комплекс оборудования, который используется для организации процесса въезда и выезда на огражденную территорию парковки или паркинга, а также для оплаты за услуги парковки. Автоматизированная парковочная система является максимально доступным и удобным вариантом для использования на предприятиях с большой нагрузкой транспортного движения. Поэтому важно с помощью вероятностных и статистических методов разработать алгоритмы, позволяющие оптимизировать работу АПС.

Целью работы является определение вероятностных и статистических методов, наиболее подходящих для разработки математической модели, позволяющей оптимизировать работу парковочной системы (на примере аэропорта Домодедово), а также решить проблемы, связанные с распознаванием номерных знаков.

2. Основная часть

Как было сформулировано выше, цель данной работы — определить наиболее подходящие для решения поставленной задачи (оптимизация работы парковочной системы) математические методы и приемы.

Для этого, в первую очередь, нужно четко сформулировать этапы решения поставленной задачи:

1. проведение анализа текущего состояния АПС, выявление основных проблем; разбиение модели АПС на составные части (система въезда/выезда, система распознавания номерного знака, система обработки поступающих запросов);
2. рассмотрение модели функционирования въезда/выезда, определение оптимальной длины подъездного пути;
3. собрать и изучить статистические данные по следующим факторам, которые могут влиять на качество распознавания номерного знака: время суток, погодные условия, расположение считывающей камеры — влияют на степень освещенности номерного знака; длина подъездного пути (острова), угол подъезда транспортного средства, угол расположения считывающей камеры — влияют на качество распознавания номерного знака;

4. вычисление средней вероятности распознавания номерного знака;
5. с помощью регрессионных моделей определение по собранной статике факторов, наиболее влияющих на качество распознавания номерных знаков;
6. определение оптимального (в том числе и с финансовой точки зрения) числа операторов парковки;
7. построение математической модели на основе скрытых цепей Маркова для более качественной обработки (разрешения) конфликтов с возможностью обучения (перенастройки);
8. определение оптимальных входных параметров для алгоритма изменения оперативного контура управления системой, обмен сигналами, программное обеспечение, какая камера должна быть, ручное управление, освещённость.

Решение первых двух задач подразумевает обладание необходимыми профессиональными компетенциями, связанными со спецификой проводимого исследования.

Для решения третьей и четвертой задач можно применять стандартные методы математической статистики и теории вероятностей [1, 2].

Пятая задача подразумевает владение аппаратом регрессионного анализа [3-6] и ее решение может быть осуществлено посредством применения функций статистического пакета R [7, 8].

Для решения шестой задачи (нахождение оптимального числа операторов) модель АПС можно представить в виде системы массового обслуживания (или сети массового обслуживания - для более сложной модели) с несколькими разнотипными входящими потоками и несколькими обслуживающими приборами. Тогда данная проблема может быть решена с помощью методов теории массового обслуживания (ТМО) [9] и теории оптимальных задач в ТМО [10].

Наконец, последние две задачи могут решаться с применением теории скрытых Марковских моделей [11], успешно реализуемых в пакете R [12, 13].

3. Заключение

Целью работы является определение вероятностных и статистических методов, наиболее подходящих для разработки математической модели, позволяющей оптимизировать работу парковочной системы (на примере аэропорта Домодедово), а также решить проблемы, связанные с распознаванием номерных знаков/

По результатам проведенных исследований (в рамках решения первых пяти сформулированных выше задач) были сделаны следующие предложения по оптимизации работы парковочных систем, направленные на повышение вероятности распознавания номерных знаков, снижение количества конфликтов и, за счет этого, снижение нагрузки с операторов парковок и повышение пропускной способности:

1. Удлинить острова, что бы исключить некорректный подъезд транспортного средства.
2. Привести освещение в соответствии со стандартом или выше на въездных и выездных группах всех парковок.
3. Заменить все камеры верхнего исполнения на камеры нижнего исполнения.

Благодарности

Публикация подготовлена при поддержке Программы РУДН «5-100».

Литература

1. Бочаров П. П., Печинкин А. В. Теория вероятностей. Математическая статистика. — М.: Гардарики, 1998.
2. Гмурман В. Е. Теория вероятностей и математическая статистика. — М.: Высшая школа, 1997.

3. Носко В. П., Эконометрика. В 2-х книгах. Кн. 1. Ч. 1. Основные понятия, элементарные методы. Ч. 2. Регрессионный анализ временных рядов. — М.: Дело, 2011.
4. Кремер Н.Ш. Эконометрика. Учебник и практикум, М.: Юрайт, 2017.
5. Айвазян С. А. Эконометрика – 2. Продвинутый курс с приложениями в финансах. — М.: Магистр, 2018.
6. Rawlings J. O., Pantula S. G., Dickey D. A. Applied Regression Analysis. A Research Tool. Second Edition, Springer, 1998.
7. Crowley M. The R book. — John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, 2007.
8. Ulrike Gromping. Inference with Linear Equality and Inequality Constraints Using R. The Package ic.infer. URL: <http://cran.gis-lab.info/web/packages/ic.infer/vignettes/ic.infer.pdf>
9. Бочаров П. П., Печинкин А. В. Теория массового обслуживания. — М.: РУДН, 1995.
10. Shaler Stidham Jr. Optimal Design of Queueing Systems. — Taylor & Francis Group, LLC, 2009.
11. Ching W., Ng M. Markov Chains: Models, Algorithms and Applications. — Shanghai, China: Springer, 2006.
12. Lin Himmelman. Package HMM. URL: <https://cran.r-project.org/web/packages/HMM/HMM.pdf>
13. Visser I., Speekenbrink M., depmixS4: An R Package for Hidden Markov Models. URL: <https://cran.r-project.org/web/packages/depmixS4/vignettes/depmixS4.pdf>

UDC 519.218, 519.22

Application of Statistical and Probabilistic Methods for Optimization of Parking Systems Operation and for the Recognition of License Plates

D. M. Ahmethanova*, I. S. Zaryadov*†

* *Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

† *Institute of Informatics Problems, FRC CSC RAS
44-2 Vavilova str., Moscow 119333, Russian Federation*

Email: 1032143094@rudn.university, zaryadov_is@rudn.university

The main probabilistic and statistical methods that can be used to solve problems of optimization of the parking system, as well as those associated with the recognition of license plates, are formulated.

Key words and phrases: automated parking system, license plate recognition, regression models, hidden Markov models, optimization problems in queuing theory.

UDC 004.4

Statistical Analysis of Inflation in Contemporary Russia

N. V. Baylo, S. I. Matyushenko, D. A. Pyatkina

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: 1032163680@rudn.university, matyushenko_si@rudn.university, pyatkina_da@rudn.university

The dependence of the growth rate of inflation in Russia on a number of macroeconomic indicators, among which the price of oil, the dollar rate, the unemployment rate, the GDP value, per capita incomes of the population, is investigated, on the basis of statistical data. The study is carried out using multidimensional statistical analysis methods using statistical packages SPSS and Eviews. As a result, the nature and strength of the relationship between the growth rate of inflation and the indicators listed above was established, the regression equation was constructed on the basis of which a forecast for the immediate future was made, the dynamics of inflation and its relationship with per capita incomes were studied.

Key words and phrases: inflation growth rate, correlation-regression analysis, regression equation, time series, crosscorrelation.

1. Introduction

Inflation is understood as a long and steady trend towards an increase in the average price level and depreciation of money. The inflationary process traditionally acts as one of the indicators of the effectiveness of the economy. In this regard, an important role is played by a clear understanding of what and to what extent it influences the growth rate of inflation and what measures must be taken to reduce it. The availability of statistical data makes it possible to use methods of multivariate statistical analysis to answer the questions posed. In the next part of the work we will use the method of correlation-regression analysis to determine the dependence of the rate on a number of macroeconomic indicators. In the third part of the work, inflation is considered as a time series. The tendency of development of inflation in modern Russia is investigated and the analysis of the dependence between inflation and incomes per capita is made taking into account the delay effect.

2. Inflation and its relationship with macroeconomic factors

Define the rate of inflation, as:

$$Y(t) = \frac{(S(t) - S(t-1))}{S(t-1)} \cdot 100, \quad t = 2, \dots, n$$

where $S(t)$ — the cost of the consumer basket in the month t according to the Federal State Statistics Service [1], $t = 1, 2, \dots, n$; $t = 1$ corresponds to January 2008, and $t = n$ — to December 2016.

As factors affecting Y , we consider the following indicators: X_1 — an increase in the value of the dollar against the ruble; X_2 — the growth rate of the price of Brent crude oil; X_3 — rate of increase in the number of unemployed; X_4 — the rate of change in GDP; X_5 — per capita monetary income of the population. These variables are not multicollinear, and their connection with Y can be characterized as moderate and positive for X_1, X_2, X_3 and as a moderate negative for X_4, X_5 :

$$Y(X) = 0,864 + 0,023 \cdot X_2 + 3,883 \cdot X_3 - 0,076 \cdot X_4 - 0,014 \cdot X_5. \quad (1)$$

Equation (1) was obtained using the statistical package SPSS [2]. From the received data it is possible to conclude the following: the equation is significant, all parameters

entering into the model (except X_1) are significant; approximately 33% of the variation in the rate of inflation is accounted for in the model and is due to the influence of the included factors. The relatively low value of the R -square coefficient indicates the presence of the delay effect, a phenomenon in which inflation changes not simultaneously with the macroeconomic indicators under consideration, but with a certain time shift. Therefore, to study the process of inflation, it is necessary to use the time series apparatus.

3. Dynamics of inflation and its relationship with per capita incomes

Using the Eviews statistical package, we constructed a time series (data from 2000 to the present time [1]) of inflation and identified its main trend using double exponential smoothing and Holt-Winters methods without seasonality (Figure 1).

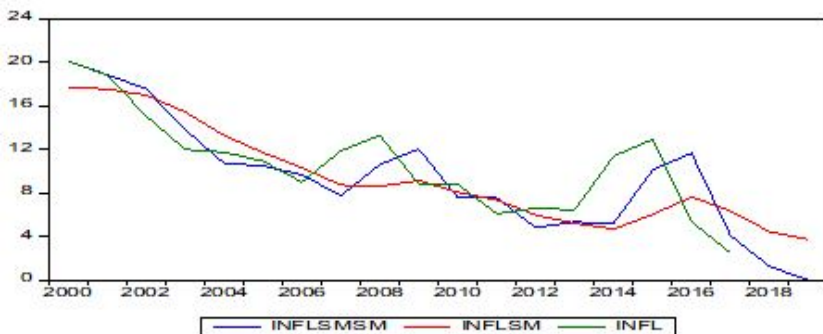


Figure 1. The trend of inflation in modern Russia

From the graphs (Figure 1), we see a tendency to reduce inflation with two upsurge in the crisis 2008 and 2014 years. In particular, according to our forecast, inflation in 2019 will be 3.4%.

Analyzing the observed trend with the help of a regression model which reflects the dependence of the inflation rate (Y) on the average per capita incomes of the population (X_5) and its lagged values, one can conclude that a decrease in household incomes entails a decrease in inflation with a delay of two months (only X_5 lagged values for the two previous months, and the coefficients at them are positive).

Also, using the crosscorrelation method, we found a positive correlation between Y and X_5 with the value in the first and third month. The absolute value of the correlation in the second month is small and lies within the statistical error.

Both methods show that a decrease in the incomes of the population entails a decrease in inflation with a delay of two months on average.

In addition, using standard time series analysis procedures [3], we constructed models for two series: inflation and population incomes in the form of ARIMA (2,0,3). The parameters of the models for both series coincided, which was another confirmation of the proximity of inflation and population incomes taking into account the delay effect.

4. Conclusion

Inflation depends on many macroeconomic indicators. However, the study of this connection requires taking into account the delay effect and should be carried out using time series and methods of cross-correlation. We were particularly interested in the problem of the dependence of inflation on income per capita. In the first part, for data obtained at the same time, we found that there is an inverse relationship between inflation and income (the corresponding coefficient in the regression equation is -0.014). However, carrying out the research taking into account the time factor, it turned out that the relationship between inflation and the population's incomes is direct and most strongly manifested with a delay of two months. In other words, two months after the falling incomes of the population, sellers, reacting to this drop, reduce or fix prices, thereby lowering the rate of inflation. This conclusion leads to the question: is today's inflation reduction the result of a skillful monetary policy of the Central Bank or is it the result of a fall in real incomes of the population?

References

1. The site of the Federal State Statistics Service - URL: <http://www.gks.ru/>
2. Multivariate statistical analysis in economic problems: computer modeling in SPSS. Textbook. allowance / under. Ed. I. V. Orlova. Moscow: University textbook, 2014. 310 p., in Russian.
3. V. N. Afanasyev, M. M. Yuzbashev. Time series analysis and forecasting. Moscow: Finance and statistics; INFRA, 2010, 320 p., in Russian.

УДК 004.4

Статистический анализ инфляции в современной России

Н. В. Байло, С. И. Матюшенко, Д. А. Пяткина

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: 1032163680@rudn.university, matyushenko_si@rudn.university, pyatkina_da@rudn.university

На основе статистических данных за период с января 2008 года по декабрь 2016 года исследуется зависимость темпа роста инфляции в России от ряда макроэкономических показателей, среди которых цена на нефть, курс доллара, уровень безработицы, величина ВВП, среднедушевые доходы населения. Исследование проводится с помощью методов многомерного статистического анализа с применением статистических пакетов SPSS и Eviews. В результате установлен характер и сила связи между темпом роста инфляции и перечисленными выше показателями, построено уравнение регрессии, на основании которого может быть составлен прогноз на ближайшую перспективу, изучена динамика инфляции и ее связь с доходами на душу населения.

Ключевые слова: темп роста инфляции, корреляционно-регрессионный анализ, уравнение регрессии, временной ряд, кросскорреляция.

УДК 519.25

Применение метода SSA к различным прикладным задачам

Д. А. Боровских*, Гбавле Люк-Дональд Дилие*, В. А. Петров†,
А. С. Савин*, В. К. Садовников*, А. А. Хохлов*

* Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198

† Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198

Email: dmirbrv@gmail.com, lucdalie@mail.ru, vladimirap@yandex.ru, assavin90@gmail.com,
sadvonikovv@yandex.ru, khokhlov_aa@rudn.university

Метод анализа временных рядов SSA (Singular Spectrum Analysis) изначально был разработан для применения в экономических задачах. Однако, впоследствии он стал применяться и для решения других задач, не связанных с экономикой. В работе описана постановка некоторых задач, для решения которых предлагается применять метод SSA.

Ключевые слова: SSA, SVD, гусеница.

1. Введение

Метод SSA успешно применяется для анализа временных рядов в течение длительного времени [1]. Суть метода заключается в том, что исходный временной ряд $f(t_i) = x_i$, где $i = 1, \dots, N$, N — целое положительное число, записывается в прямоугольную траекторную матрицу A , первая строка которой состоит из M первых членов временного ряда: $A[1] = x_1, x_2, \dots, x_M$, вторая строка начинается со 2-го члена ряда: $A[2] = x_2, x_3, \dots, x_{M+1}$ и так далее. Число M называется длиной гусеницы.

На следующем этапе траекторная матрица умножается на себя же транспонированную с некоторым коэффициентом, затем, при необходимости, полученная квадратная матрица нормируется и центрируется (в зависимости от решаемой задачи). Далее полученная матрица \tilde{A} при помощи одного из методов численной диагонализации матриц (например, метода Якоби), приводится к виду: $\tilde{A} = S^{-1} * D * S$, где S — матрица собственных векторов матрицы \tilde{A} , а D — диагональная матрица собственных значений матрицы \tilde{A} .

Далее проводится процедура анализа полученных собственных значений. Каждое из них соответствует тренду, влияющему на поведение исходного временного ряда. Например, выбрав некоторые из значений и проведя процедуру ганкеллизации, можно восстановить отфильтрованный временной ряд, выяснив их влияние на исследуемый процесс.

Также существуют модификации метода SSA для работы с двухмерными и многомерными временными рядами (CSSA — Coupled SSA и MSSA — Multi SSA).

2. Задача быстрого сравнения изображений

Предположим, у нас имеется графическое изображение в формате Bitmap. Если представить цветное значение каждого пикселя как элемент временного ряда, то к такому изображению можно применить метод SSA. Выделив главные компоненты ряда, можно сформировать «эталон» изображения. Далее, если требуется принять решение относительно другого изображения, насколько оно близко к эталонному, то при помощи того или иного инструмента анализа главных компонент ряда, можно принять такое решение с той или иной точностью.

Такая методика может быть использована на практике в случаях, когда подключение нейронной сети для распознавания изображения неоправданно дорого с точки зрения вычислительных и финансовых ресурсов. Например, алгоритм может быть использован для быстрого распознавания номеров на автомобилях при проезде контрольно-пропускных пунктов.

3. Задача классификации активности человека

Также при помощи метода SSA можно анализировать показания акселерометров носимых мобильных устройств и делать выводы о качестве и характере движений человека, например, такой задачей является анализ качества выполнения различных спортивных упражнений. Проанализировав показания акселерометров, закрепленных на кистях и лодыжках, можно, например, определить, насколько точно человек выполняет те или иные упражнения. Это актуально для восточных единоборств, художественной гимнастики, танцев и других видов спорта, требующих точного выполнения тех или иных движений.

Примерный алгоритм анализа будет следующий:

1. Показания акселерометра считываются в процессе выполнения упражнения с носимого устройства и записываются в локальную базу данных.
2. Из выборки данных акселерометра строится трехмерный временной ряд (так как имеются данные по ускорениям по трем осям).
3. Временной ряд анализируется методом SSA [1] (при помощи троекратного применения одномерного алгоритма SSA, либо при помощи применения MSSA).
4. На основании анализа тем или иным образом делается вывод о характере движения, либо его качестве.

4. Задача анализа показаний биоимпедансометра

Анализ импеданса тела (БИА) представляет собой один из методов классификации состава тела и классификации значений ткани тела. Метод БИА использует ток небольшой частоты, приложенный к телу пациента. В соответствии с этим измеряется напряжение и электрические свойства тела используются для оценки состава тела. Этот метод очень популярен, поскольку он неинвазивный, быстрый, портативный и недорогой.

Суть метода заключается в том, что через тело пациента пропускают высокочастотный 1 кГц до 1 МГц слабый ток, используя два или более подключений электродов, а затем измеряют амплитуду тока после прохождения через тело человека.

Исходя из полученных данных, а также антропометрических исходных данных: возраст, пол, расовая и этническая принадлежность, линейные и весовые параметры тела (длина, масса, объем тела), на основании имеющихся статистических данных делается вывод о составе тела человека (соотношение мышечной ткани, жировой ткани, костной ткани, воды и других составляющих) и другие выводы (в зависимости от условий и целей).

Авторами предлагается использовать алгоритм SSA следующим образом. Результаты измерений амплитуды прошедшего через тело человека тока записываются в виде временных рядов, и далее анализируются с помощью SSA (и его модификации для многомерных временных рядов — MSSA). Вычисляются собственные значения и собственные векторы траекторной матрицы и производится непосредственно анализ функции, представленной собственными значениями.

Исходя из предположения, что при выборе параметров SSA (окно Гусеницы), траекторная матрица содержит в себе вполне исчерпывающую информацию об исходном временном ряде, анализ спектра матрицы позволит делать различные выводы о причинах поведения исходного временного ряда — то есть об искомых для исследователя параметрах.

5. Заключение

В работе рассмотрены несколько задач, для решения которых авторами предлагается использовать метод SSA. Авторами были проведены серии вычислительных экспериментов, которые показали уместность применения методики SSA к решению описанных задач.

Литература

1. Данилов Д. Л., Жиглявский А. А. Главные компоненты временных рядов: метод «Гусеница». — Санкт-Петербургский Государственный Университет, 1997.

UDC 519.25

Different Applications of SSA Method

D. A. Borovskikh*, Gbavle Luk-Donald Dilie*, V. A. Petrov[†], A. S. Savin*,
V. K. Sadovnikov*, A. A. Khokhlov*

* *Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

[†] *Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: dmirbv@gmail.com, lucdalie@mail.ru, vladimirap@yandex.ru, assavin90@gmail.com,
sadovnikovv@yandex.ru, khokhlov_aa@rudn.university

SSA (Singular Spectrum Analysis) method was originally developed for time series analysis in problems of economics. However, this method can be applied for solving problems in a large number of other areas. This paper describes formulations if several problems for which SSA method can be applied.

Key words and phrases: SSA, svd.

УДК 519.863

Информационно-математическая модель поддержки принятия решений по выводу из эксплуатации объектов использования атомной энергии

В. В. Бочкарёв*, А. В. Крянев^{†‡},
Д. С. Смирнов[§], Р. Б. Шарафутдинов*

* НТЦ ядерной и радиационной безопасности
Малая Красносельская ул., д. 2/8, к. 5., Москва, Россия, 107140

[†] Кафедра прикладной математики
Национальный исследовательский ядерный университет «МИФИ»
Каширское ш., д. 31., Москва, Россия, 115409

[‡] Объединенный институт ядерных исследований
ул. Жолио-Кюри, д. 6, Дубна, Московская обл., Россия, 141980

[§] Отделение социально-экономических наук офиса образовательных программ
Национальный исследовательский ядерный университет «МИФИ»
Каширское ш., д. 31., Москва, Россия, 115409

Email: avkryanev@mephi.ru

В докладе представлена структура информационно-математической модели поддержки принятия решений по выводу из эксплуатации объектов использования атомной энергии. В предлагаемой структуре предусмотрено наличие схемы анализа различных вариантов вывода из эксплуатации объектов использования атомной энергии, основанной на многокритериальном подходе. В частности, наряду с обычными выводами из эксплуатации (до уровня «зеленой лужайки» и захоронением радиоактивных отходов на месте) объекта использования атомной энергии в модели предусмотрено их оптимальное сочетание. В модели учитывается неопределенности в исходных данных, в том числе, данных, определяемых экспертно, неопределенность которых описывается с помощью нечетких множеств. Информационно-математическая модель оценивает выводимый объект по большому числу показателей, сводя их к нескольким агрегированным критериям, из которых наиболее значимыми являются сводный показатель потенциальной опасности и сводный финансовый показатель с учетом уровней их неопределенности. Схема позволяет выбрать оптимальный вариант вывода из эксплуатации объекта использования атомной энергии, соответствующего оптимальному соотношению критериев, характеризующих обеспечение значений показателей ядерной-радиационной безопасности и экономических показателей выводимого из эксплуатации объекта использования атомной энергии.

Ключевые слова: информационно-математическая модель (ИММ), вывод из эксплуатации (ВЭ), объект использования атомной энергии (ОИАЭ), многокритериальный подход, сводный комплексный критерий.

1. Введение

Создание ИММ ВЭ ОИАЭ является актуальной для органов управления использованием атомной энергией, при разработке и актуализации концепций ВЭ ОИАЭ, их соответствующему обоснованию и представлению в орган государственного регулирования материалов, обосновывающих безопасность и экономическую целесообразность деятельности по реализации ВЭ ОИАЭ.

Задачи выбора варианта реализации ВЭ ОИАЭ является задачей многокритериального выбора на конечном множестве альтернатив, осложненная: неопределенными значениями оценок объемов производимых работ; размытостью используемых мнений и оценок экспертов; интервальной оценкой последствий реализации того или иного варианта реализации ВЭ ОИАЭ; необходимостью учета возможных рисков и неопределенностей; многовариантными подходами к обращению с РАО, образующимися при демонтаже и дезактивации основного технологического оборудования и инженерной инфраструктуры.

2. Содержание математической модели

Содержание части математической модели, связанной с оптимизацией очередности ВЭ ОИАЭ, зависит от определения агрегированного показателя, числовое значение которого определяется по частным показателям для каждого выводимого из эксплуатации ОИАЭ. Таким единственным итоговым показателем и для отдельного ОИАЭ, и для всей совокупности потенциальных для ВЭ ОИАЭ, может являться сводный показатель потенциальной опасности R , вобравший в себя остальные частные показатели, кроме показателя экономической эффективности. Этот показатель является расчетной величиной, учитывающей, в том числе, и оценки экспертного характера и, тем самым, содержит в себе различный уровень неопределенности для различных ОИАЭ. Показатель R позволяет определять его для совокупности ОИАЭ простым суммированием этого показателя для ОИАЭ, входящих в рассматриваемую совокупность.

В ИММ предусмотрен учет всех необходимых физических и финансово-экономических показателей для каждого варианта по ВЭ ОИАЭ, включая обеспечения ЯРБ в процессе захоронения на месте и транспортировки тех частей РАО, которые будут вывезены, переработаны и захоронены в пунктах захоронения и экономических показателей всех затрат, связанных с захоронением, транспортировкой и их переработкой.

Каждый из вариантов по ВЭ ОИАЭ будет оцениваться тремя группами показателей: качественных, физических и финансово-экономических. Аналитический блок ИММ ВЭ обеспечит свертку показателей по каждой группе и даст итоговую оценку варианта ВЭ, используя многокритериальный подход и степень неопределенности каждого показателя.

В предлагаемой ИММ ВЭ наряду с двумя традиционными вариантами ВЭ: 1) ВЭ до уровня «Зеленой лужайки», 2) ВЭ с захоронением на месте выводимого из эксплуатации объекта, будет возможность просчитывать различные другие варианты реализации ВЭ, когда часть РАО будет захоронена на месте, а другая часть вывозится для захоронения в других местах переработки и долговременного захоронения.

ИММ ВЭ ОИАЭ позволит: получать исходные показатели реализации ВЭ ОИАЭ с помощью предварительных расчетов или с помощью экспертных оценок с учетом неопределенности их значений; оценивать варианты по ВЭ ОИАЭ, сравнивать их и выбирать наиболее предпочтительный вариант на основе объективных качественных, физических и финансово-экономических показателей; проводить анализ вариантов реализации ВЭ ОИАЭ методами нечеткой логики и нечетких множеств; выбирать оптимальный вариант реализации ВЭ каждого рассматриваемого выводимого из эксплуатации ОИАЭ с учетом степени обеспечения ЯРБ и финансово-экономических показателей на транспортировку, переработку и захоронение РАО; решать задачи оптимизации реализации ВЭ совокупности ОИАЭ при любых заданиях и описании неопределенностей и риска.

ИММ ВЭ ОИАЭ может быть использована для: оценки вариантов по ВЭ ОИАЭ; сопоставления вариантов реализации ВЭ ОИАЭ по большому набору показателей, носящих как количественный, так и качественный характер, и задаваемых в процессе решения задачи; выбора наиболее предпочтительного варианта реализации ВЭ ОИАЭ, в том числе, в условиях неопределенности и риска и с учетом финансово-экономических показателей; детального обоснования выбора варианта реализации ВЭ, как отдельного ОИАЭ, так и совокупности ОИАЭ

3. Заключение

В докладе предложено содержание ИММ поддержки принятия решений по ВЭ ОИАЭ. Модель позволяет оценивать и сравнивать между собой несколько ВЭ ОИАЭ с учетом качественных, физических и финансово-экономических показателей каждого ОИАЭ и сравнивать между собой различные варианты вывода из эксплуатации ОИАЭ.

Литература

1. Experience Using Phenomena Identification and Ranking Technique (PIRT) for Nuclear Analysis. Brookhaven National Laboratory. BNL-76750-2006-CP. 2006.
2. Wilson G. E., Boyack B. E. The Role of the PIRT Process in Experiments, Code Development and Code Applications Associated with Reactor Safety Analysis // Nuclear Engineering and Design. — November 1998. — Vol. 186 (1–2).
3. Микони С. В. Многокритериальный выбор на конечном множестве альтернатив. — СПб: Издательство «Лань», 2009.
4. Атенков А. В., Зарубин В. С., Канатников А. М. Введение в методы оптимизации. — М.: Финансы и статистика, 2008.
5. Ногин В. Д. Принятие решений при многих критериях. — М.: СПб: Издательство «ЮТАС», 2007.
6. Selection of decommissioning. Strategies: Issues and factors. International Atomic Energy Agency. IAEA-TECDOC-1478. 2005.
7. Decommissioning Nuclear Power Plants: Policies, Strategies and Costs. OECD Nuclear Energy Agency, ISBN 92-64-10431-3. 2003.

UDC 519.863

Information and Mathematical Model for Supporting Decision-Making on the Decommissioning of Nuclear Power Facilities

V. V. Bochkaev*, A. V. Kryanev^{†‡}, D. S. Smirnov[§], R. B. Sharafutdinov*

* *Scientific and Engineering Centre for Nuclear and Radiation Safety
Malaya Krasnoselskaya st. 2/8, bld. 5, Moscow, 107140, Russian Federation*

[†] *Department of Applied Mathematics
National Research Nuclear University MEPhI
Kashirskoe shosse, 31, Moscow, 115409, Russian Federation*

[‡] *Joint Institute for Nuclear Research
Joliot-Curie 6, Dubna, Moscow region, 141980, Russian Federation*

[§] *Office of Academic Programs: Division of Economic and Social Sciences
National Research Nuclear University MEPhI
Kashirskoe shosse, 31, Moscow, 115409, Russian Federation*

Email: avkryanev@mephi.ru

The report proposes the structure of the information-mathematical model for supporting decision-making on the decommissioning of nuclear power facilities. The proposed structure provides for the existence of a scheme for analyzing the various options for decommissioning the use of atomic energy based on a multicriteria approach. In particular, along with the usual inferences from operation (to the level of the "green lawn" and the disposal of radioactive waste on site), the object of using atomic energy in the model provides for their optimal combination. The model takes into account the uncertainties in the initial data, including data determined expertly, whose uncertainty is described by fuzzy sets. The information-mathematical model evaluates the output object for a large number of indicators, reducing them to several aggregated criteria, of which the most significant are the composite indicator of potential hazard and the consolidated financial indicator, taking into account the levels of their uncertainty. The scheme allows you to choose the optimal option for decommissioning an object using atomic energy that corresponds to the optimal correlation of criteria characterizing the maintenance of values of nuclear and radiation safety indicators and economic indicators of the decommissioned object of nuclear energy use.

Key words and phrases: information-mathematical model (IMM), decommissioning (DE), object of atomic energy use (OAEU), multicriteria approach, consolidated complex criterion.

UDC 517.972

On Poisson Brackets in Spaces of B_u -potentials

S. A. Budochkina

*Mathematical Institute n.a. S.M. Nikol'skii
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: budochkina_sa@rudn.university

A relationship between B_u -Hamiltonian operators and Poisson brackets is established.

Key words and phrases: Poisson brackets, B_u -potentials, Jacobi identity.

1. Introduction

The theory of Hamiltonian systems has very efficient methods for the integration and qualitative investigation of differential equations [1]. However, many important systems of equations cannot be represented in the form of canonical Hamiltonian equations. The analysis of such equations has led to the development of a nonclassical Hamiltonian formalism, which is based on the notion of noncanonical Poisson brackets [2, 3].

The problem of the introduction of Poisson brackets in spaces of B_u -potentials is of special interest. The use of the operator approach allows to introduce Poisson brackets for different types of operators (ordinary differential operators, differential operators with partial derivatives, integro-differential operators, etc.) and in spaces of Euler and non-Euler functionals.

In what follows, we will use the notation and terminology of [4–6].

2. B_u -Hamiltonian operators and Poisson brackets

Consider a nonlocal bilinear form

$$\Phi_1(\cdot, \cdot) \equiv \langle \cdot, \cdot \rangle : V_1 \times V_1 \rightarrow \mathbb{R} \quad (1)$$

and an operator $B_u : V_1 \rightarrow V_1$.

Definition. A linear operator $G_u : D(G_u) \subset V_1 \rightarrow V_1$ is called B_u -Hamiltonian (with respect to bilinear form (1)), if $\forall h, v, g \in V_1$ it satisfies the following conditions:

$$\langle g, B_u G_u h \rangle = - \langle h, B_u G_u g \rangle,$$

$$\begin{aligned} \langle v, B_u G'_u(g; G_u h) + B'_u(G_u h; G_u g) \rangle + \langle g, B_u G'_u(h; G_u v) + B'_u(G_u v; G_u h) \rangle + \\ + \langle h, B_u G'_u(v; G_u g) + B'_u(G_u g; G_u v) \rangle = 0. \end{aligned}$$

Denote by $\mathcal{F}_{\Phi_1, B_u}$ a real linear space of functionals $F[t, \cdot] : U_1 \subseteq V_1 \rightarrow \mathbb{R}$ with representation

$$\delta F[t, u, h] = \langle (B_u - grad)_{\Phi_1} F[t, u], B_u h \rangle \quad \forall u \in U_1, \quad \forall h \in D(F'_u, B_u).$$

For any elements $F_1, F_2 \in \mathcal{F}_{\Phi_1, B_u}$ define their product by the formula

$$\{F_1, F_2\}[t, u] = \langle K_1(t, u), B_u G_u K_2(t, u) \rangle \quad \forall u \in U_1, \quad (2)$$

where $K_1(t, u) = (B_u - grad)_{\Phi_1} F_1[t, u]$, $K_2(t, u) = (B_u - grad)_{\Phi_1} F_2[t, u]$.

Theorem 1. If the operator G_u is B_u -Hamiltonian with respect to symmetric nondegenerate bilinear form $\Phi_1(1)$, B_u is an invertible operator, then $\forall F_1, F_2, F_3 \in \mathcal{F}_{\Phi_1, B_u}$ and $\forall \lambda_1, \lambda_2 \in \mathbb{R}$ the following conditions are fulfilled on U_1 :

$$\begin{aligned} \{F_1, F_2\} &= -\{F_2, F_1\}, \\ \{\lambda_1 F_1 + \lambda_2 F_2, F_3\} &= \lambda_1 \{F_1, F_3\} + \lambda_2 \{F_2, F_3\}, \\ \{F_1, \lambda_1 F_2 + \lambda_2 F_3\} &= \lambda_1 \{F_1, F_2\} + \lambda_2 \{F_1, F_3\}, \\ \{F_1, \{F_2, F_3\}\} + \{F_2, \{F_3, F_1\}\} + \{F_3, \{F_1, F_2\}\} &= 0 \quad (\text{Jacobi identity}). \end{aligned}$$

Theorem 2. If the operator G_u is B_u -Hamiltonian with respect to symmetric nondegenerate bilinear form $\Phi_1(1)$, B_u is an invertible operator, then product (2) is a Poisson bracket.

3. Conclusions

Using the operator approach we introduce Poisson brackets in spaces of B_u -potentials.

Acknowledgments

The work is supported by RFBR grants No 16-01-00450a and No 16-08-00558a.

References

1. V. V. Kozlov, Symmetries, topology and resonances in Hamiltonian mechanics, Izhevsk: Udmurt. Univ., 1995.
2. V. E. Zakharov, L. D. Faddeev, Korteweg-de Vries equation: A completely integrable Hamiltonian system, Functional Analysis and Its Applications (1971). 5 (4) 280–287. URL: <http://doi.org/10.1007/BF01086739>
3. L. A. Takhtadzhyan, L. D. Faddeev, The Hamiltonian approach in soliton theory, Moscow: Nauka, 1986.
4. V. M. Savchin, Mathematical methods in mechanics of infinite-dimensional nonpotential systems, Moscow: PFU, 1991.
5. S. A. Budochkina, V. M. Savchin, On B_u -Hamiltonian equations in mechanics of infinite-dimensional systems, Doklady Mathematics (2011). 84 (1) 525–526. URL: <http://doi.org/10.1134/S1064562411050012>
6. S. A. Budochkina, On a representation of an operator equation with first time derivative in the form of a B_u -Hamiltonian equation, Differential Equations (2013). 49 (2) 176–186. URL: <http://doi.org/10.1134/S0012266113020043>

УДК 517.972

О скобках Пуассона в пространствах B_u -потенциалов

С. А. Будочкина

*Математический институт им. С.М. Никольского
Российский университет дружбы народов
ул. Миклухо-Маклая, д. 6, Москва, Россия, 117198*

Email: budochkina_sa@rudn.university

Установлена взаимосвязь между B_u -гамильтоновыми операторами и скобками Пуассона.

Ключевые слова: скобки Пуассона, B_u -потенциалы, тождество Якоби.

UDC 004.4

On the Volatility Function in a Modified Dispersion Model with Constant Elasticity (CEV)

A. E. Buuruldai

*Nicol'skii Mathematical Institute
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: aivo4ka@mail.ru

The problem of definition the function of volatility in the modified dispersion model with constant elasticity (CEV) is considered in paper. An algorithm was constructed to find the volatility function from a given density of distribution the underlying asset for solving the problem of definition the function of volatility, which resulted in a formula that allows calculating the volatility of the European call option for any given distribution without knowing its value. The obtained result is used to construct the volatility function in the CEV model.

Key words and phrases: volatility, European call option, density distribution, differential equation and CEV model.

1. Introduction

The constant elasticity of variance model (CEV) is based on the assumption that risk is a neutral process that describes the behavior of the stock price S , has the form

$$dS_t = (r - q)S_t dt + \delta S_t^{\beta+1} dB_t, \quad t \geq 0, \quad S_0 = S > 0,$$

where r — risk-free interest rate and let $\mu = r - q$, dB_t — Wiener process, δ — volatility parameter and β — constant [1].

Formulas for calculating the value of European call options in the CEV model are as follows:

$$c(S; K, T) = \begin{cases} e^{-qT} SQ(\xi; n - 2, y_0) - e^{-rT} K(1 - Q(y_0; n, \xi)), & \beta > 0, \\ e^{-qT} SQ(y_0; n, \xi) - e^{-rT} K(1 - Q(\xi; n - 2, y_0)), & \beta < 0, \end{cases}$$

where

$$n = 2 + \frac{1}{|\beta|}, \quad \xi = \frac{2\mu S^{-2\beta}}{\delta^2 \beta (e^{2\mu\beta T} - 1)}, \quad y_0 = \frac{2\mu K^{-2\beta}}{\delta^2 \beta (1 - e^{-2\mu\beta T})},$$

$Q(x, y, z)$ — the noncentral chi-squared distribution with the parameter of noncentricity of z and y degrees of freedom is less than the number x . Let us consider the case when $\beta = -\frac{1}{2}$, then we obtain a stochastic differential equation of the form

$$dS_t = \mu S_t dt + \delta S_t^{\frac{1}{2}} dB_t, \quad t \geq 0, \quad S_0 = S > 0.$$

Density distribution

$$\rho_\mu(T; S; S_T) = e^{-\mu T} \frac{2S^{\frac{1}{2}} S_T^{-\frac{1}{2}} e^{\frac{\mu T}{2}}}{\delta^2 T} \exp\left(-\frac{2(S + S_T e^{-\mu T})}{\delta^2 T}\right) \cdot I_\alpha\left(\frac{4S^{\frac{1}{2}} S_T^{\frac{1}{2}} e^{-\frac{\mu T}{2}}}{\delta^2 T}\right),$$

where $I_\alpha(z)$ — Bessel function,

$$I_\alpha(z) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m! \Gamma(m + \alpha + 1)} \left(\frac{x}{2}\right)^{2m + \alpha}, \quad \alpha > 0.$$

The Bessel function for $\alpha = 1$ can be represented in the form $I(z) = \frac{x}{4} [1 + O(x^2)]$:
 $x \rightarrow 0$

Suppose $\delta = 1$, then the distribution density

$$\rho_\mu(T; S; K) = 2e^{-\mu T} \frac{S}{T^2} \exp\left(-\frac{2(S + Ke^{-\mu T})}{T}\right) = 2e^{-\mu T} \frac{S}{T^2} \cdot e^{-\frac{2S}{T}} \cdot e^{-\frac{2Ke^{-\mu T}}{T}}.$$

2. Volatility in the CEV model

In the value of the European options $c(T, K)$ is known for all maturities T and all price – strikes K , then for the volatility function σ the Dupire formula [2]:

$$\sigma(T, K) = \sqrt{\frac{\frac{\partial c}{\partial T} + r \cdot K \cdot \frac{\partial c}{\partial K}}{\frac{1}{2} K^2 \frac{\partial^2 c}{\partial K^2}}} \tag{1}$$

in which to find the value of the volatility function the variables T, K are replaced by the variables t, S . Then, using formula (1), we can calculate the volatility of the European call and using the homotopy formula [3]:

$$c(K, T, S, T) = 2e^{-\mu T} \frac{S}{T^2} \cdot e^{-\frac{2S}{T}} \int_0^1 (1 - \tau) e^{-\frac{2K\tau e^{-\mu T}}{T}} K^2 d\tau + \varphi_1 K + \varphi_0,$$

where $\mu = r - q$, q — dividends. In the calculations dividends, were not taken into account $q = 0$

$$\frac{\partial c}{\partial K} = -\frac{S}{T} e^{-\frac{2Ke^{-\mu T}}{T} + \frac{2S}{T}}, \quad \frac{\partial^2 c}{\partial K^2} = 2e^{-\mu T} \frac{S}{T^2} \cdot e^{-\frac{2S}{T} - \frac{2Ke^{-\mu T}}{T}}.$$

So we get:

$$\sigma^2(T, K) = \frac{(2S + \mu T^2) \left(-e^{-\frac{2Ke^{-\mu T}}{T}} + 1\right)}{2K^2 S e^{-2\mu T}} + \frac{1}{K S e^{-\mu T}},$$

where $r = 0$

$$\sigma^2(T, K) = \frac{1}{K^2 S} \left(S \left(1 - e^{-\frac{2K}{T}}\right) + K\right).$$

3. Conclusions

Finding the volatility of function from a given density of the distribution underlying asset a volatility function which based on the constructed algorithm is found in the modified variance model with constant elasticity (CEV).

References

1. D. Davydov, V. Linetsky. Pricing and Hedging Path-Dependent Options Under the CEV Process, 2001, URL: <https://doi.org/10.1287/mnsc.47.7.949.9804>
2. Bruno Dupire. Pricing with a Smile, 1994.
3. A. D. Bosov, Yu. N. Orlov, The non-stationary time series forecasting with the use of Fokker–Plank equation and evolution equations for momentum of empirical distribution function, Keldysh Institute preprints, 2013, 003, 30 p.

УДК 004.4

О функции волатильности в модифицированной модели дисперсии с постоянной эластичностью (CEV)

А. Э. Буурулдай

*Математический институт им С. М. Никольского
Российский университет дружбы народов
ул. Миклухо-Маклая, д. 6, Москва, Россия, 117198*

Email: aivo4ka@mail.ru

Рассматривается задача нахождения функции волатильности в модифицированной модели дисперсии с постоянной эластичностью (CEV). Для решения поставленной задачи был построен алгоритм нахождения функции волатильности по заданной плотности распределения базового актива, в результате которого получена формула, которая позволяет вычислить волатильность европейского опциона колл для любого заданного распределения, не зная его стоимость. Полученный результат применен для построения функции волатильности в CEV модели.

Ключевые слова: волатильность, европейский колл опцион, плотность распределение, дифференциальное уравнение и CEV модель.

УДК 519.624.3

Моделирование дальнего действия тепловых изменений металла, облучаемого нанокластерами меди

Б. Батгэрэл[‡], С. Н. Димова[†], И. В. Пузынин^{*}, Т. П. Пузынина^{*},
З. К. Тухлиев^{*}, И. Г. Христов[†], Р. Д. Христова[†], З. А. Шарипов^{*}

** Лаборатория информационных технологий
Объединённый институт ядерных исследований
ул. Жолио-Кюри, д. 6, Дубна, Московская область, Россия, 141980*
† Софийский университет «Св. Климент Охридски», София, Болгария
*‡ Монгольский государственный университет науки и технологии
Улан-Батор, Монголия*

Email: zafar@jinr.ru, zarif@jinr.ru

При облучении нанокластерами металлической мишени в малом объеме взаимодействия происходит выделение большого количества энергии, что приводит к структурным изменениям в металле. Источником структурных изменений могут быть тепловые процессы или упругие взаимодействия (расталкивание атомов мишени). Экспериментальные исследования фиксируют результаты этих структурных изменений в мишени. Применение методов математического моделирования позволяет «увидеть» динамику процесса облучения, включая тепловые и структурные изменения. Поэтому математическое моделирование остается важным инструментом изучения процессов облучения мишеней нанокластерми.

В предыдущих работах [1, 2] авторами был обнаружен эффект волнового переноса тепла при моделировании термических процессов в металлах, облучаемых нанокластерами меди с энергией в диапазоне 10-50 эв/атом. В настоящей работе, используя ранее полученный результат, проведено моделирование эффекта дальнего действия методом молекулярной динамики. Эффект дальнего действия заключается в изменении структуры и свойств облучаемой мишени на глубинах, значительно превышающих область первичного выделения энергии, при взаимодействии энергетических потоков с поверхностью твердых тел. В данной работе показано, что могут возникнуть структурные изменения кристаллической решетки на глубине мишени, превышающей глубину проникновения нанокластеров.

Ключевые слова: молекулярная динамика, эффект переноса тепла, эффект дальнего действия, нанокластеры.

1. Введение

Облучение металлических мишеней импульсными пучками ионов можно приближенно моделировать как облучение последовательностью нанокластеров, и в этом случае задачу можно рассматривать в рамках метода молекулярной динамики. Действительно, непосредственное применение молекулярной динамики для задачи облучения тяжелыми ионами высоких энергий невозможно, что связано с неупругими взаимодействиями при высоких энергиях налетающего тяжелого иона. Однако, облучение нанокластерами металлических мишеней [1, 2] сходно с моделированием облучения другими источниками: тяжелыми ионами высоких энергий и импульсными пучками. При облучении тяжелыми ионами высоких энергий в локальной области мишени также выделяется большое количество энергии в единице объема. При моделировании в рамках молекулярной динамики [3] это имеет большое значение.

В настоящее время для применения методов молекулярной динамики существуют готовые программные пакеты (LAMMPS, DL-POLY, NAMD и т.д.). В данной работе моделирование проводилось с помощью пакета LAMMPS [4], установленного на кластере HybriLit [5], который модифицировался в соответствии с задачей моделирования одновременного взаимодействия нескольких нанокластеров с мишенью.

Моделирование процесса взаимодействия нанокластеров с металлической мишенью проводилось следующим образом. В рамках молекулярно-динамического подхода решаем задачу облучения металлической мишени несколькими нанокластерами с различной энергией. Далее с помощью полученного решения вычисляем температурные профили мишени в зависимости от времени. Для этого материал мишени разбиваем на ячейки и в каждой ячейке вычисляем температуру для каждого шага по времени. При этом надо учитывать, что количество частиц в ячейке должно быть больше 100. Таким образом, мы получаем температурное поле. В рамках молекулярной динамики расчетной областью являлся параллелепипед со сторонами $30 \times 30 \times 16$ нм. В качестве межатомного потенциала использовался универсальный ЕАМ — потенциал для меди, встроенный в пакет LAMMPS [4].

Настоящая работа отличается от предыдущих работ [1, 2] тем, что металлическая мишень облучается не одним нанокластером а одновременно двумя нанокластерами, области взаимодействия которых с поверхностью мишени находятся друг от друга на определенном расстоянии (2–3 нм). Так как волновой перенос тепла распространяется по диагонали, это приводит к тому, что в глубине мишени происходит слияние движущихся областей с высокой температурой. Температура в области слияния резко возрастает, превышая температуру плавления мишени, что может привести к структурным изменениям мишени.

На рис. 1 приведены результаты при одновременном облучении двумя нанокластерами.

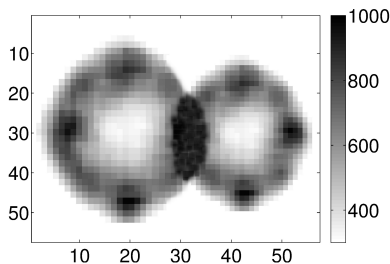


Рис. 1. Температурное поле в мишени, на глубине 5 нм при облучении нанокластерами

Поскольку глубина проникновения нанокластера составляет 2–3 нм, результаты свидетельствуют об эффекте дальнего действия тепловых изменений металла.

Благодарности

Работа выполнена при финансовой поддержке гранта РФФИ № 17-01-00661-а и при частичной поддержке гранта Полномочного представителя Республики Болгария в ОИЯИ.

Литература

1. Batgerel B., Dimova S., Puzynin I. et al. // EPJ Web Conf. — 2018. — Vol. 173. — P. 06001.
2. Batgerel B., Didyk A. Yu., Puzynin I. V. et al. // Journal of Surface Investigation. — 2015. — Vol. 9. — Pp. 1026–1030.
3. Холмуродов Х. Т., Алтайский М. В., Пузынин И. В. и др. // ЭЧАЯ. — 2003. — Т. 34, вып. 2. — С. 472.

4. Plimpton S. Fast Parallel Algorithms for Short-Range Molecular Dynamics // J. Comp. Phys. — 1995. — Vol. 117. — Pp. 1–19. doi:10.1006/jcph.1995.1039.
5. HybriLit. URL: <http://hybrilit.jinr.ru/>

UDC 519.624.3

Modeling of Long-Range Thermal Changes of Metal Irradiated by Copper Nanoclusters

**B. Batgerel[‡], S. N. Dimova[†], I. V. Puzynin^{*}, T. P. Puzynina^{*},
Z. K. Tukhliev^{*}, I. G. Hristov[†], R. D. Hristova[†], Z. A. Sharipov^{*}**

** Laboratory of Information Technologies*

Joint Institute for Nuclear Research

6, Joliot-Curie str., Dubna, Moscow Region, 141980, Russian Federation

[†] Sofia University "St. Kliment Ohridski", Sofia, Bulgaria

*[‡] Mongolian State University of Science and Technology
Ulaanbaatar, Mongolia*

Email: zafar@jinr.ru, zarif@jinr.ru

When nanoclusters interact with a metal target, a large amount of energy is realized in a small volume which leads to structural changes. The changes can be a consequence of thermal processes or of elastic interactions (repulsion of target atoms). The experimental research fixes the results of these structural changes in the target. The application of methods of mathematical modeling makes possibility to "see" the dynamics of the irradiation process, including thermal and structural changes. Therefore, the mathematical modeling remains an important tool for such studies.

In previous works [1, 2], the authors discovered the effect of wave heat transfer in the simulation of thermal processes in metals irradiated by copper nanoclusters with energy in the range 10-50 eV/atom. In this work, using the previously obtained result, thermal long-range effect is modeled by the molecular dynamics method. The long-range effect consists in changing the structure and properties of the irradiated target at depths significantly exceeding the region of primary energy flows interacting with the surface of target. In this paper it is shown that possible structural changes in the crystal lattice at the target depth exceeding the penetration depth of nanoclusters can occur.

Key words and phrases: molecular dynamics, heat transfer effect, long-range effect, nanoclusters.

Определение коэффициентов гармонической линеаризации для детерминистической нелинейной системы с управлением

Т. Р. Велиева, А. В. Завозина, А. В. Королькова

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: velieva_tr@rudn.university, 1032171772@rudn.university, akorolkova@sci.pfu.edu.ru

При исследовании систем с управлением было выявлено явление глобальной синхронизации узлов в сети, которая проявляется в виде автоколебательного режима. Наличие данного режима ухудшает характеристики сети, например, пропускную способность. Для определения параметров возникновения автоколебательного режима в нашем научном коллективе было решено применить метод гармонической линеаризации.

Ключевые слова: активное управление трафиком, автоколебательный режим, RED, метод гармонической линеаризации.

1. Введение

Для решения задачи нахождения параметров автоколебаний важно использовать метод, который учитывает все характеристики исходной системы. Поэтому предложено использовать метод гармонической линеаризации, отличительной чертой которого является зависимость результирующей системы от амплитуд и частот периодических процессов, что позволяет исследовать свойства нелинейных систем.

2. Гармоническая линеаризация нелинейного элемента системы

Для описания системы с управлением будем использовать следующую непрерывную модель взаимодействия входящего потока типа TCP и маршрутизатора, обрабатывающего поток по алгоритму типа RED (Random Early Detection) [1], аналогично работам [2, 3]:

$$\begin{cases} \dot{W}(t) = \frac{1}{T(Q, t)} - \frac{W(t)W(t - T(Q, t))}{2T(t - T(Q, t))} p(t - T(Q, t)); \\ \dot{Q}(t) = \frac{W(t)}{T(Q, t)} N - C; \\ \dot{\hat{Q}}(t) = -w_q C \hat{Q}(t) + w_q C Q(t). \end{cases}$$

Здесь используются следующие обозначения: $W(t)$ — средний размер окна TCP; $Q(t)$ — среднее значение размера очереди; $\hat{Q}(t)$ — экспоненциально-взвешенное скользящее среднее (EWMA) среднего размера очереди; C — интенсивность обслуживания очереди; T — полное время двойного оборота; $T = T_p + Q/C$, где T_p — время двойного оборота свободной сети (без учёта задержек в оборудовании); Q/C — время нахождения пакета в очереди; N — количество сессий TCP; $p(t)$ — функция сброса пакетов.

Прежде чем применять метод гармонической линеаризации, необходимо линеаризовать функцию сброса пакетов для алгоритма RED [4]. Линеаризованная функция сброса пакетов P_{RED} выглядит следующим образом:

$$P_{RED} = \begin{cases} 0, & 0 < \hat{Q} \leq Q_{\min}, \\ \frac{p_{\max}}{Q_{\max} - Q_{\min}}, & Q_{\min} < \hat{Q} \leq Q_{\max}, \\ 0, & \hat{Q} > Q_{\max}. \end{cases}$$

Здесь Q_{\min} и Q_{\max} — пороговые значения средневзвешенного значения длины очереди, p_{\max} — максимальный уровень сброса пакетов.

Вычислим коэффициенты гармонической линеаризации $\varkappa_0(A, \omega, x_0)$, $\varkappa(A, \omega, x_0)$ и $\varkappa'(A, \omega, x_0)$ для статической нелинейности P_{RED} :

$$\begin{aligned} \varkappa_0(A, \omega, x_0) &= \frac{1}{2\pi} \int_0^{2\pi} P_{RED}(x_0 + A \sin(\omega t)) d(\omega t); \\ \varkappa(A, \omega, x_0) &= \frac{1}{A\pi} \int_0^{2\pi} P_{RED}(x_0 + A \sin(\omega t)) \sin(\omega t) d(\omega t); \\ \varkappa'(A, \omega, x_0) &= \frac{1}{A\pi} \int_0^{2\pi} P_{RED}(x_0 + A \sin(\omega t)) \cos(\omega t) d(\omega t). \end{aligned} \tag{1}$$

В зависимости от соотношений между порогами Q_{\min} , Q_{\max} , сдвигом x_0 и амплитудой A возможно получить разные пределы интегрирования. Продемонстрируем случай (см. рис.1), когда

$$Q_{\min} < x_0 < Q_{\max}, \quad x_0 - A < Q_{\min}, \quad x_0 + A > Q_{\max}. \tag{2}$$

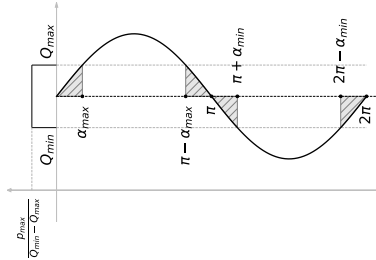


Рис. 1. Определение пределов интегрирования

Найдём значения пределов интегрирования α_{\min} и α_{\max} :

$$\begin{aligned} x_0 + A \sin \alpha_{\min} &= Q_{\min}, \quad \sin \alpha_{\min} = \frac{Q_{\min} - x_0}{A}; \quad \cos \alpha_{\min} = \sqrt{1 - \frac{(Q_{\min} - x_0)^2}{A^2}}; \\ x_0 + A \sin \alpha_{\max} &= Q_{\max}, \quad \sin \alpha_{\max} = \frac{Q_{\max} - x_0}{A}; \quad \cos \alpha_{\max} = \sqrt{1 - \frac{(Q_{\max} - x_0)^2}{A^2}}. \end{aligned}$$

Таким образом, проинтегрировав (1) в области (2), получим:

$$\begin{aligned} \varkappa_0(A, x_0) &= \frac{p_{\max}}{2\pi(Q_{\max} - Q_{\min})} \left[2 \arcsin \left(\frac{Q_{\max} - x_0}{A} \right) + 2 \arcsin \left(\frac{Q_{\min} - x_0}{A} \right) \right]; \\ \varkappa(A, x_0) &= -\frac{2p_{\max}}{\pi A(Q_{\max} - Q_{\min})} \sqrt{1 - \frac{(Q_{\max} - x_0)^2}{A^2}}; \\ \varkappa'(A, x_0) &= 0. \end{aligned}$$

3. Заключение

Полученные аналитические выражения могут быть использованы для реализации программы в системе компьютерной алгебры для последующего исследования исходной системы для определения областей возникновения автоколебаний.

Благодарности

Работа частично поддержана грантом РФФИ № 16-07-00556. Также публикация подготовлена при поддержке программы РУДН «5-100».

Литература

1. Floyd S., Jacobson V., Random Early Detection Gateways for Congestion Avoidance // IEEE/ACM Transactions on Networking. — 1993. — Vol. 1, no. 4. — Pp. 397–413.
2. Misra V., Gong W. B., Towsley D., Fluid-Based Analysis of a Network of AQM Routers Supporting TCP Flows with an Application to RED // ACM SIGCOMM Computer Communication Review. — 2000. — Vol. 30, No 4. — Pp. 151–160.
3. Велиева Т. Р., Королькова А. В., Кулябов Д. С., Сантуш Б. А., Модель управления очередями на маршрутизаторах // Вестник РУДН. Серия: Математика. Информатика. Физика. — 2014. — № 2. — С. 81–92.
4. Кулябов Д. С., Королькова А. В., Велиева Т. Р., Применение метода гармонической линеаризации к исследованию автоколебательного режима систем с управлением // Вестник РУДН. Серия: Математика. Информатика. Физика». — 2017. — Т. 25, № 3. — С. 234–252.

The Determination of the Coefficients of Harmonic Linearization for Deterministic Nonlinear System with Control

T. R. Velieva, A. V. Zavozina, A. V. Korolkova

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia ((RUDN University))
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: velieva_tr@rudn.university, 1032171772@rudn.university, akorolkova@sci.pfu.edu.ru

The study of systems with management revealed the phenomenon of global synchronization of nodes in the network, which manifested in the form of a self-oscillating mode. The existence of this mode impares the characteristics of the network, for example, bandwidth. Our research team have decided to apply the method of harmonic linearization for determination of the parameters of the self-oscillating regime.

Key words and phrases: traffic active management, self-oscillating mode, RED, method of harmonic linearization.

УДК 004.021:519.2:519.6

NS-2 как средство верификации амплитудно-частотных характеристик системы с управлением

Т. Р. Велиева, А. В. Завозина, А. В. Королькова

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: velieva_tr@rudn.university, 1032171772@rudn.university, korolkova_av@rudn.university

В работе предложено использовать средство имитационного моделирования NS-2 для верификации амплитудно-частотных характеристик, полученных в аналитической модели модуля управления трафиком.

Ключевые слова: активное управление трафиком, имитационное моделирование, NS2, автоколебательный режим, RED.

1. Введение

Исследуется явление возникновения автоколебательного режима в системах с активным управлением очередями на маршрутизаторах в сетях передачи данных. В основу построения аналитической модели системы с управлением, в которой возможно возникновение автоколебательного режима, положена модель, предложенная в [1]. Задача выявления причин возникновения автоколебаний динамических переменных в подобных моделях рассмотрена в работах [2, 3]. В этих работах предложена методика получения амплитудно-частотных характеристик динамических переменных рассматриваемой модели.

После проведения численного анализа возникла задача верификации полученных результатов. Одним из способов верификации является проведение имитационного эксперимента. Поскольку в качестве примера системы с управлением рассматривается модель модуля активного управления трафиком в сетях передачи данных, то предлагается реализовать имитационную модель в средстве имитационного моделирования сетевых протоколов NS-2.

2. Имитационное моделирование на основе NS2

Пакет NS2 [4, 5] является средством имитационного моделирования сетевых протоколов. Данный пакет можно использовать как альтернативу натурному эксперименту.

В качестве основного языка разработки в NS-2 применяется язык TCL [6, 7]. При проведении имитационного эксперимента использовалась так называемая гантельная топология (см. рис. 1). Стандартным образом были определены в модели TCP-источники и TCP-приёмники, маршрутизаторы с управлением по алгоритму RED (Random Early Detection) и соединения между элементами сети. Дополнительные сессии TCP имитировались путём добавления дополнительных источников. Также были заданы значения параметров, влияющие на возникновение автоколебательного режима — пороговые значения алгоритма управления RED, размер очереди на маршрутизаторе, размер пакета, максимальная вероятность сброса пакетов в алгоритме RED.

Результатом проведения имитационного эксперимента является большой объем данных trace-файла, полностью определяющий поведение объектов системы. После дополнительной обработки данных из trace-файла можно, используя реализацию на языке Julia быстрого преобразования Фурье, построить спектральный портрет автоколебательного режима. При проведении анализа данных была использована

система сборки *Snakemake* [<https://snakemake.readthedocs.io>]. В качестве результата получен, в частности, амплитудный спектр для среднего размера окна TCP (рис. 2) изучаемой модели.

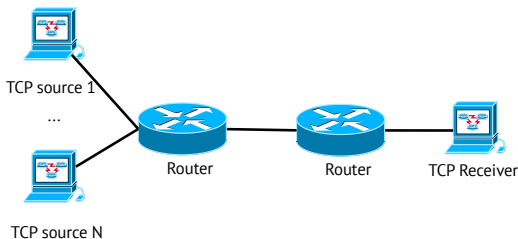


Рис. 1. Гантельная топология

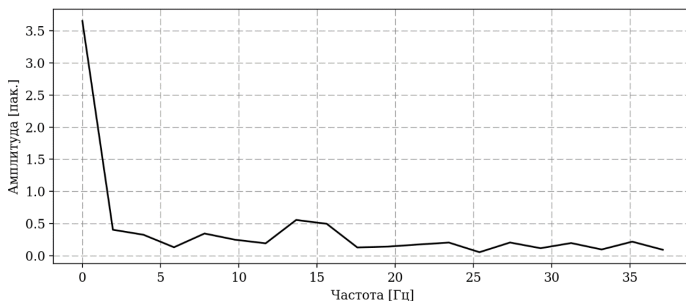


Рис. 2. Амплитудный спектр для среднего размера окна TCP

Полученный таким образом на базе имитационного моделирования амплитудно-частотный спектр динамических переменных системы (размера окна и/или очереди) можно сравнить с аналогичным спектром, полученным при аналитических вычислениях, и таким образом верифицировать данные исследования.

3. Заключение

В ходе исследования разработан комплекс программ для проведения имитационного эксперимента по исследованию автоколебательного режима в системах с управлением и верификации теоретических результатов.

Благодарности

Работа частично поддержана грантом РФФИ № 16-01-20379. Также публикация подготовлена при поддержке программы РУДН «5-100».

Литература

1. Misra V., Gong W.-B., Towsley D. Stochastic Differential Equation Modeling and Analysis of TCP-window-size Behavior // Proceedings of IFIP WG 7.3 Performance. — 1999. — Vol. 99. — Access mode: <http://dna-pubs.cs.columbia.edu/citation/paperfile/24/Misra99-TCP-Stochastic.pdf>.
2. Кулябов Д. С., Королькова А. В., Велиева Т. Р. Применение метода гармонической линеаризации к исследованию автоколебательного режима систем с управлением // Вестник РУДН. Серия «Математика. Информатика. Физика». — 2017. — Т. 25, № 3. — С. 234–252. — Режим доступа: <http://journals.rudn.ru/miph/article/view/16206>.
3. Kulyabov D. S., Korolkova A. V., Velieva T. R., Eferina E. G., Sevastianov L. A. The Methodology of Studying of Active Traffic Management Module Self-oscillation Regime // DepCoS-RELCOMEX 2017. Advances in Intelligent Systems and Computing / Ed. by W. Zamojski, J. Mazurkiewicz, J. Sugier, T. Walkowiak, J. Kacprzyk. — Cham : Springer International Publishing, 2018. — Vol. 582 of Advances in Intelligent Systems and Computing. — P. 215–224. — Access mode: http://link.springer.com/10.1007/978-3-319-59415-6_21.
4. Issariyakul T., Hossain E. Introduction to Network Simulator NS2. — 2012. — Vol. 978146141414. — P. 510.
5. Altman E., Jiménez T. NS Simulator for Beginners // Synthesis Lectures on Communication Networks. — 2012. — Vol. 5, no. 1. — P. 1–184.
6. Welch B., Jones K. Practical Programming in Tcl and Tk. — 4th edition. — Prentice Hall, 2003. — P. 960.
7. Nadkarni A. P. The Tcl Programming Language: A Comprehensive Guide. — CreateSpace Independent Publishing Platform, 2017. — P. 668.

UDC 004.021:519.2:519.6

NS-2 as a Verification Means of the Amplitude-Frequency Characteristics of the Control System

T. R. Velieva, A. V. Zavozina, A. V. Korolkova

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: velieva_tr@rudn.university, 1032171772@rudn.university, korolkova_av@rudn.university

It was suggested to use the simulation tool NS-2 for verification of amplitude-frequency characteristics obtained in the analytical model of the traffic management module.

Key words and phrases: active queue management, simulation, NS-2, self-oscillating, RED.

УДК 519.218.21

Моделирование распространения сетевых червей с помощью стохастических дифференциальных уравнений

М. Н. Геворкян, А. В. Демидова, Т. С. Демидова, А. А. Соболев

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: gevorkyan_mn@rudn.university, demidova_av@rudn.university, 1032152607@pfur.ru, 1032152613@pfur.ru

В данной работе рассматривается модель SIR, используемая для описания динамики распространения сетевых червей. Целью данной работы является применение стохастических эпидемиологических моделей к моделированию динамики распространения компьютерных вирусов локальной сети. Учет вероятностных явлений при моделировании сложных систем может улучшить математическую модель исследуемой системы.

Ключевые слова: модель эпидемии, SIR, сетевые черви, математическое моделирование.

1. Введение

Современный мир невозможно представить без компьютеров, которые стали частью нашей повседневной жизни. Данные устройства, являясь частью компьютерных сетей, используются для хранения конфиденциальной информации. Однако вся эта информация находится под угрозой сетевых вирусов. Одной из разновидностей сетевых вирусов являются вредоносные программы, способные к самостоятельному поиску новых узлов для заражения и использующие для своего распространения локальные и глобальные компьютерные сети, создавая свои копии. Современные средства защиты не всегда могут оперативно реагировать на эпидемии сетевых червей, поэтому актуальной является задача создания средств, способных предотвратить или сдержать эпидемию на ранних стадиях.

2. Модель эпидемии SIR

В данной работе используется модель SIR, описывающая затухание сетевых эпидемий. В этой модели узлы могут быть в одном из трех состояний: S — уязвимое, I — зараженное, R — невосприимчивое. В данной системе узлы не появляются и не выходят, т.е. система замкнутая. Для нее верно утверждение $S + I + R = N$. Вводим переменные $s = S/N$, $i = I/N$ и $r = R/N$ для описания относительных численностей узлов. Предполагаем, что заражение всех узлов происходит равновероятно. Тогда, данная модель может быть описана системой дифференциальных уравнений: [1, 3]

$$\begin{cases} \frac{ds}{dt} = -\beta is, \\ \frac{di}{dt} = \beta is - \gamma i, \\ \frac{dr}{dt} = -\gamma i. \end{cases} \quad (1)$$

В системе (1) коэффициент β — скорость заражения узлов, γ — скорость выздоровления узлов.

3. Самосогласованная стохастическая модель

Метод построения самосогласованных стохастических моделей [5, 6] базируется на комбинаторной методологии, которая описана в [7] и [8]. Метод заключается в построении схемы взаимодействия. Это символическая запись возможных взаимодействий элементов системы. Далее записывается интенсивность переходов и основное кинетическое уравнение. Для него записывается приближенное уравнение Фоккера–Планка и эквивалентное ему стохастическое дифференциальное уравнение в форме Ланжевена. Состояние системы описывается вектором $x = s, i, r$. Схема взаимодействия и вектор R имеют вид:

$$S + I \xrightarrow{\beta} 2I, \quad R^1 = (-1, 1, 0), \quad I + A \xrightarrow{\gamma} R, \quad R^2 = (0, -1, 1). \quad (2)$$

В схеме взаимодействия (2) первая строка описывает зараженный узел с уязвимым состоянием, в результате чего в системе образуется новый зараженный узел. Вторая строка описывает применение к зараженному узлу «лекарства», вследствие чего данный узел переходит в невосприимчивое к заражению состояние. Уравнение Фоккера–Планка записывается следующим образом:

$$\frac{\partial P(s, i, r)}{\partial t} = - \sum_a [\partial A_a(s, i, r) P(s, i, r, t)] + \frac{1}{2} \sum_{a,b} \partial_a \partial_b [B_{ab}(s, i, r) P(s, i, r, t)],$$

где вектор сносов A и матрица диффузии B имеют вид:

$$A = \begin{pmatrix} -\beta is \\ \beta is - \gamma i \\ \gamma i \end{pmatrix}, \quad B = \begin{pmatrix} \beta is & -\beta is & 0 \\ -\beta is & \beta is + \gamma i & -\gamma i \\ 0 & -\gamma i & \gamma i \end{pmatrix}.$$

Эквивалентное уравнению Фоккер–Планка уравнение в форме Ланжевена записывается как:

$$dx(t) = a(x, t)dt + b(x, t)dW,$$

где $x(t) \in R^3$ — функция состояния системы, а $W \in R^3$ — стандартный трёхмерный винеровский процесс, который реализуется как $dW = \varepsilon \sqrt{dt}$, где $\varepsilon \sim N(0, 1)$ — нормальное распределение со средним, равным 0, и дисперсией, равной 1. Для коэффициентов справедливо следующее соотношение:

$$a(x) = A(x), \quad b(x) = B(x)B(x)^T.$$

4. Численный анализ динамики распространения сетевых червей

Численный анализ динамики распространения сетевых червей проводился с помощью разработанного комплекса программ, реализованного на языке Python 3 с использованием библиотек NumPy и SciPy, а также Matplotlib для графического представления вычислений [9, 10]. Кроме того, реализована возможность интерактивного подбора параметров модели с помощью библиотеки Ipywidgets. Численный эксперимент проводился для сети из $N = 1000$ узлов. Количество инфицированных узлов в начальный момент времени и параметры γ и β подбирались интерактивно. На рис. 1 представлены результаты численного эксперимента со значениями $I(0) = 8$ узлов, $\gamma = 0,36$, $\beta = 0,05$.

Для стохастической модели численный эксперимент показал практически полное совпадение траекторий для стохастического и детерминистического случаев. Таким образом изучение стохастической модели распространения сетевых червей показало, что в некоторых случаях введение стохастики слабо влияет на поведение системы, поэтому для изучения системы можно рассматривать ее детерминистическое приближение, которое определяется матрицей сносов $A(x)$.

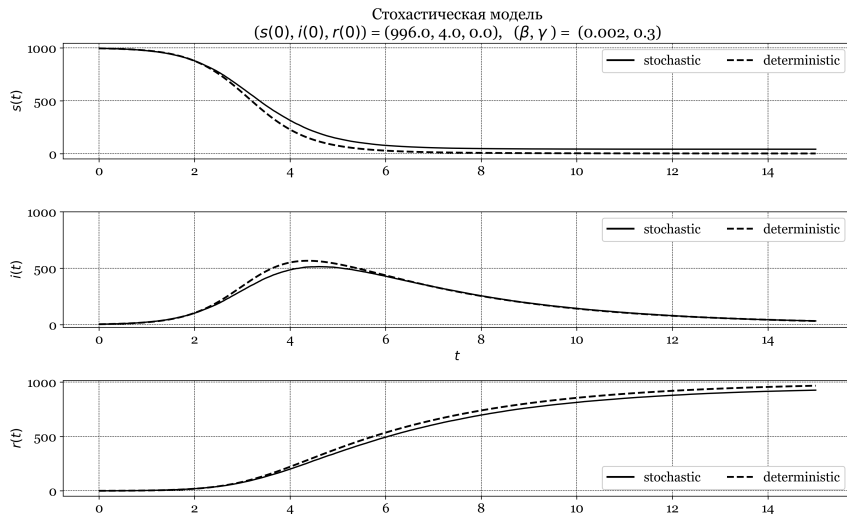


Рис. 1. Динамика распространения сетевых червей

5. Заключение

В работе рассмотрены детерминистическая и стохастическая модели распространения сетевых червей. Описана структура стохастических моделей, выписано уравнение Фоккера—Планка, сформулировано правило перехода к стохастическому дифференциальному уравнению в форме Ланжевена. Проведен сравнительный анализ детерминистической и стохастической моделей, показавший, что в данной модели не учитываются многие факторы, влияющие на динамику распространения сетевых червей.

Литература

1. Котенко И. В., Воронцов В. В. Аналитические модели распространения сетевых червей // Труды СПИИРАН, 2007.
2. Захарченко А. Черводинамика: причины и следствия // Защита информации. Конфидент. — 2004. — № 2. — С. 50–55.
3. Herbert W. Hethcote The Mathematics of Infectious Diseases // SIAM Review. — 2000. — Vol. 42 (4). — Pp. 599–653. doi:10.1137/S0036144500371907.
4. Harko T., Lobo F. S. N., and Mak M. K. Exact analytical solutions of the Susceptible-Infected-Recovered (SIR) epidemic model and of the SIR model with equal death and birth rates // Applied Mathematics and Computation. — 2014. — Vol. 236. — Pp. 184–194.
5. Кулябов Д. С., Демидова А. В. Введение согласованного стохастического члена в уравнение модели роста популяций // Вестник РУДН. Серия: Математика. Информатика. Физика. — 2012. — № 3. — С. 69–78.
6. Demidova A. V., Gevorkyan M. N., Egorov A. D., Korolkova A. V., Kulyabov D. S., Sevastyanov L. A. Influence of stochasticization to one-step model // Вестник РУДН. Серия: Математика. Информатика. Физика. — 2014. — № 1. — С. 71–85.

7. Van Kampen N.G. Stochastic Processes in Physics and Chemistry. — Elsevier Science, 1992.
8. Gardiner C.W. Handbook of Stochastic Methods: for Physics, Chemistry and the Natural Sciences (Springer Series in Synergetics, 1985).
9. Gevorkyan M.N., Velieva T.R., Korolkova A.V., Kulyabov D.S., Sevastyanov L.A. Stochastic Runge–Kutta Software Package for Stochastic Differential Equations // In Dependability Engineering and Complex Systems. — Springer International Publishing, 2016. — Pp. 169–179.
10. Eferina E.G., Korolkova A.V., Gevorkyan M.N., Kulyabov D.S., Sevastyanov L.A. One-Step Stochastic Processes Simulation Software Package // Вестник РУДН. Серия: Математика. Информатика. Физика. —2014. — № 3. — С. 46–59.

UDC 519.218.21

Simulation of the Spread of Networks Worms with Using Stochastic Differential Equations

M. N. Gevorkyan, A. V. Demidova, T. S. Demidova, A. A. Sobolev

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: gevorkyan_mn@rudn.university, demidova_av@rudn.university, 1032152607@pfur.ru,
1032152618@pfur.ru

In this paper, we consider the SIR model used to describe the dynamics of the spread of network worms. The aim of this work is to apply stochastic epidemiological models to modeling the dynamics of computer viruses spreading on the local network. Taking into account probabilistic phenomena in modeling complex systems can improve the mathematical model of the studing system.

Key words and phrases: epidemic model, SIR, network worms, mathematical modeling.

УДК 004.4

Проблемы реализации стохастических численных методов Рунге–Кутты

М. Н. Геворкян*, А. В. Демидова*, А. В. Королькова*,
Д. С. Кулябов*, И. М. Гостев†

* *Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

† *Национальный исследовательский университет «Высшая школа экономики»
ул. Мясницкая, д. 20, Москва, Россия, 101000*

Email: gevorokyan_mn@rudn.university

В данной работе рассматриваются стохастические численные методы типа Рунге–Кутты с слабой и сильной сходимостями для систем стохастических дифференциальных уравнений в форме Ито. Дается краткий обзор основных стохастических численных методов и сведений из теории стохастических дифференциальных уравнений. Далее излагается и мотивируется подход к реализации данных методов с помощью генерации исходного кода. Обсуждаются детали реализации и используемые языки программирования и библиотеки.

Ключевые слова: стохастические дифференциальные уравнения; стохастические численные методы; автоматическая генерация кода; Python; Julia; шаблонизатор.

1. Введение

В статье авторов [1] описывалась реализация стохастических численных методов типа Рунге–Кутты на языке Python с использованием библиотек NumPy и SciPy. Выбор языка был продиктован удобством программирования и возможностью работать с многомерными массивами как с тензорами. Однако быстрейшее создание функций была на низком уровне. В данной работе мы описываем альтернативный подход к реализации стохастических численных методов, основанный на генерации кода численных методов.

2. Анализ трудностей реализации стохастических численных методов Рунге–Кутты

Стохастические методы Рунге–Кутты значительно сложнее своих классических аналогов. Можно указать следующие факторы, усложняющие реализацию стохастических методов в программном виде.

- Общая громоздкость формул численных схем.
- При выборе конкретного метода надо учитывать какой тип сходимости необходимо обеспечить для данной конкретной задачи, а также какое из стохастических уравнений необходимо решать — в форме Ито или в форме Стратоновича. Это увеличивает количество алгоритмов, которые нужно запрограммировать.
- Для методов с сильной сходимостью большей единицы на каждом шаге необходимо решать ресурсоёмкую задачу по аппроксимации двукратных стохастических интегралов.
- В численной схеме присутствуют не только матрицы и векторы, но и тензоры (четырёхмерные массивы), с которыми необходимо совершать операцию свертки по нескольким индексам. Реализация свертки через суммирование с помощью обычных циклов приводит к существенному падению производительности.
- Для использования слабых методов необходимо применять метод Монте-Карло, проводя несколько серий по множеству испытаний в каждой. Так как метод Монте-Карло сходится приблизительно как $1/\sqrt{N}$, где N — число испытаний, то для достижения точности хотя бы 10^{-3} необходимо провести минимум 10^6 испытаний.

Наиболее существенное падение производительности происходит при реализации универсального алгоритма, то есть такой программы, которая может произвести расчет, используя произвольную таблицу коэффициентов. В этом случае приходится использовать большое количество вложенных циклов для того, чтобы организовать суммирование. Наличие в схемах с СДУ двойных сумм и сложной комбинации индексов в множителях, находящихся под знаком этих сумм, еще более усложняет задачу, и число вложенных циклов вырастает до шести.

Кроме этих специфических особенностей стоит упомянуть также несколько причин падения производительности, присущих также и классическим схемам, которые в стохастическом случае также играют роль. Очевидным способом хранения коэффициентов методов является использование массивов. Однако у явных методов, которые мы рассматриваем, матрица является ниже-диагональной и хранение ее в виде двумерного массива приводит к тому, что более чем половина выделенной для массива памяти тратится на хранение нулей.

Если изучить исходные коды популярных подпрограмм, реализующих классические явные вложенные методы Рунге–Кутты, то можно обнаружить, что в этих программах для хранения коэффициентов метода используется набор именованных констант, а не массивов. Это вызвано также и тем, что операции со скалярными величинами в большинстве языков программирования проводятся быстрее, чем операции с массивами.

При сохранении требования универсальности создаваемого кода и вместе с тем желание увеличить скорость вычислений и уменьшить расход памяти, привели нас к решению использовать автоматическую генерацию кода по одному шаблону для каждого отдельного метода.

Кроме выигрыша в производительности автоматическая генерация кода позволяет добавлять или изменять все функции за раз путем редактирования одного лишь шаблона, а не каждой отдельной функции. Это позволяет как уменьшить количество ошибок, так и генерировать различные варианты функций для разных целей.

3. Краткое описание автоматической генерации кода

В качестве языка для написания генератора кода используется язык Python. Исходный код и документация созданных в нашем коллективе программ доступны по ссылке bitbucket.org/mngev/sde_num_generation. Данный репозиторий содержит модуль `stochastic`, в котором реализован винеровский случайный процесс и численные методы типа Рунге–Кутты для скалярных и многомерных стохастических дифференциальных уравнений в форме Ито. Большая часть кода для этого модуля генерируется набором скриптов из директории `generator`. В ручную написан класс, реализующий многомерный винеровский процесс и методы Эйлера–Марумы.

Для работы генератора кода мы использовали библиотеку для обработки шаблонов Jinja2 [2]. Данный шаблонизатор (template engine) разрабатывался изначально для генерации HTML страниц, однако он обладает очень гибким синтаксисом и может использоваться как универсальное средство для генерации текстовых файлов любого вида, в том числе и исходных кодов на любых языках программирования. Кроме Jinja2 мы также использовали библиотеку NumPy для работы с массивами и ускорения вычислений.

Шаблоны для генерации кода представляют собой файлы с исходным кодом на языке Python со вставками специальных команд шаблонизатора Jinja2. Информация о коэффициентах методов хранится отдельно в структурированном виде в формате JSON. Это позволяет легко добавлять новые методы и изменять старые путем редактирования и JSON файлов. В настоящее время генерируются методы с коэффициентами, представленными в работах [3–5].

В качестве языка для уже сгенерированных функций используется сам python с активным применением библиотеки NumPy, которая позволяет получить приемлемую производительность. Однако генерируемый код легко можно переформатировать так, чтобы он соответствовал синтаксису любого другого языка программирования.

4. Заключение

В данной работе рассмотрены стохастические численные схемы с порядком сходимости выше 0,5. Показано, что такие методы существенно сложнее эквивалентных численных методов для систем обыкновенных дифференциальных уравнений. Выделены их особенности, делающие эффективную программную реализацию таких методов не тривиальной задачей. В заключительной части статьи обсуждается подход, основанный на автоматической генерации программного кода, который позволяет получить эффективную реализацию рассматриваемых методов, давая возможность использовать любую таблицу коэффициентов. Дано краткое описание программы, созданной авторами, и ссылка на репозиторий с исходными кодами.

Благодарности

Публикация подготовлена при поддержке программы РУДН «5-100» и при финансовой поддержке РФФИ в рамках научных проектов РФФИ № 15-07-08795, 16-07-00556.

Литература

1. Gevorkyan M. N., Velieva T. R., Korolkova A. V., Kulyabov D. S., Sevastyanov L. A. Stochastic Runge–Kutta Software Package for Stochastic Differential Equations // *Dependability Engineering and Complex Systems*. — Springer International Publishing, 2016. — Vol. 470. — P. 169–179. — 1606.06604.
2. Jinja2 official site. — Access mode: <http://http://jinja.pocoo.org>.
3. Debrabant K., Rößler A. Continuous weak approximation for stochastic differential equations // *Journal of Computational and Applied Mathematics*. — 2008. — no. 214. — P. 259–273.
4. Debrabant K., Rößler A. Classification of Stochastic Runge–Kutta Methods for the Weak Approximation of Stochastic Differential Equations / *Technische Universität Darmstadt, Fachbereich Mathematik*. — 2013. — Mar. — arXiv:1303.4510v1.
5. Kloeden P. E., Platen E. *Numerical Solution of Stochastic Differential Equations*. — 2 edition. — Berlin Heidelberg New York : Springer, 1995. — ISBN: 3-540-54062-8.

UDC 004.4

Problems in Realization of Stochastic Numerical Runge–Kutta Methods

M. N. Gevorkyan*, A. V. Demidova*, A. V. Korolkova*, D. S. Kulyabov*,
I. M. Gostev†

* *Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

† *National Research University Higher School of Economics
20 Myasnitskaya str., Moscow, 101000, Russian Federation*

Email: gevorkyan_mn@rudn.university

This article discusses stochastic numerical methods of Runge-Kutta type with weak and strong convergences for systems of stochastic differential equations in Itô form. At the beginning we give a brief overview of the stochastic numerical methods and information from the theory of stochastic differential equations. Then we motivate the approach to the implementation of these methods using source code generation. We discuss the implementation details and the used programming languages and libraries

Key words and phrases: stochastic differential equations; stochastic numerical methods; automatic code generation; Python; Julia; the template engine.

УДК 519.612

Об одном итерационном методе для решения плохо обусловленных систем линейных алгебраических уравнений с разреженными матрицами

С. Ю. Гоголева, Л. Г. Зайнетдинова

*Самарский национальный исследовательский университет
им. академика С. П. Королева
ул. Московское шоссе 34А, Самара, Россия, 443086*

Email: gogoleva_s@mail.ru, zainetdinova@mail.ru

Предлагается новый подход к предобуславливанию для решения плохо обусловленных систем линейных алгебраических уравнений итерационными методами, который позволяет находить более точное решение и минимизировать заполнение разреженных матриц.

Ключевые слова: предобуславливатель, итерационные методы, разреженные матрицы, заполнение, системы уравнений.

1. Введение

Для решения систем линейных алгебраических уравнений (СЛАУ) с несимметричными плохо обусловленными разреженными матрицами больших размерностей наиболее эффективными и устойчивыми среди итерационных методов решения таких систем уравнений считаются проекционные методы. Они не требуют нахождения оптимальных итерационных параметров и обязательной априорной информации о спектре исходной матрицы, а так же допускают работу с предобуславливателями разных типов [4].

В данной работе предлагается предобуславливатель, применяемый к СЛАУ при ее решении итерационными методами подпространства Крылова и основанный на LU -разложении матрицы. Он обеспечивает меньшее заполнение матрицы, а так же более высокую скорость сходимости и точность решения по сравнению с другими предобуславливателями.

2. Предобуславливатель на основе LU -разложения с частичным выбором ведущего элемента

Рассмотрим систему алгебраических уравнений

$$Ax = b, \quad (1)$$

где матрица $A \in R^{n \times n}$, $b \in R^n$, $\det A \neq 0$. Рассмотрим вместо исходной системы (1) новую систему

$$\check{A}\check{x} = \check{b}, \quad (2)$$

где $\check{b} = M_1^{-1}b$, $\check{A} = M_1^{-1}AM_2^{-1}$, $\check{x} = M_2x$, M_1, M_2 – невырожденные матрицы размерности n . Процесс перехода от (1) к (2) с целью улучшения характеристик матрицы для ускорения сходимости к решению называется предобуславливанием, матрица M^{-1} матрицей предобуславливателя [1, 4]. Рассмотрим левое предобуславливание, которое чаще всего применяется на практике. В этом случае система (1) примет вид:

$$M^{-1}Ax = M^{-1}b, \quad (3)$$

которая в силу невырожденности M имеет то же точное решение \check{x} . Введем обозначения $\hat{A} = M^{-1}A$, $\hat{b} = M^{-1}b$, тогда (3) примет вид:

$$\hat{A}x = \hat{b}.$$

Рассмотрим предобуславливатель, применяемый к СЛАУ (1) при ее решении итерационными методами подпространства Крылова и основанный на LU -разложении матрицы A [2]. Выбор в качестве матрицы предобуславливателя $M = LU$ для СЛАУ с разреженными матрицами больших размерностей влечет за собой заполнение портрета матриц L и U . Избежать этого возможно, если в качестве предобуславливателя использовать LU -разложение матрицы A без заполнения. В этом случае матрица задается в виде $M = \bar{L}\bar{U}$, где \bar{L} и \bar{U} нижне- и верхнетреугольные матрицы, соответственно. Матрицы \bar{L} и \bar{U} вычисляются аналогично матрицам L и U , при помощи LU -разложения матрицы, лишь с тем отличием, что если $a_{ij} = 0$, то полагаем $l_{ij} = 0$ и $u_{ij} = 0$. Обозначим этот предобуславливатель $LU(0)$. Он эффективно справляется с решением задач, если матрица системы с диагональным преобладанием, но довольно грубо аппроксимирует сильно несимметричные матрицы, вследствие чего возрастает относительная погрешность решения.

Предложим способ $LU(t)$ -разложения с выбором ведущего элемента, который решает проблему деления на ноль, что позволяет применять его к невырожденным матрицам A с произвольным портретом. Использование же порога заполнения портрета матрицы сокращает объем памяти требуемый для хранения матрицы предобуславливателя. Как правило, частичного выбора ведущего элемента для построения такого предобуславливателя оказывается вполне достаточно.

Поскольку полное разложение приводит к заполнению портрета матрицы, без каких либо ограничений, прибегнем к введению некоторого уровня заполнения. В алгоритм LU -разложения с частичным выбором ведущего элемента введем параметр t , называемый критерием заполнения, и определим его следующим способом:

$$t = \|A\|_2 \varepsilon_{\text{маш}}, \quad (4)$$

где $\varepsilon_{\text{маш}}$ – машинное epsilon, а $\|A\|_2$ – евклидова норма матрицы.

Такой способ исключения ненулевых элементов плох в случае не отмасштабированных матриц, в силу большого значения $\|A\|_2$, критерий заполнения, заданный по (4), становится слишком высоким и не может быть применен. Тогда параметр задается как $t = \varepsilon$, где ε – некоторая заданная точность.

Исследование построенных алгоритмов проводилось на реальных задачах, в качестве тестовых матриц были выбраны матрицы из коллекции Harwell-Boeing [3]. Результаты заполнения матриц в результате решения СЛАУ с использованием предобуславливателя на основе LU -разложения с выбором ведущего элемента и параметра представлены в табл. 1.

Таблица 1

Заполнение портрета матриц

Название матрицы	LU-разложение	LU(t) с выбором вед. эл.
FS 183 1	12904	6128
FS 541 4	111871	39923
MCCA	1603	3735
MCFE	56073	62949
WATT 1	206542	187387

Приведем результаты решения СЛАУ методом квазимиимальных невязок для одной из матриц коллекции [3] в табл. 2.

Таблица 2

Результаты работы метода квазиминимальных невязок для матрицы FS 541 4

Предобуславливатель	Число итераций	Отн. погрешность
Отсутствует	1706	0,45e-5
Якоби	209	0,85e-5
LU-разложение	2	0,11e-10
LU(0)-разложение	9	0,11e-5
LU(t) с выбором вед. эл.	2	0,64e-11

3. Заключение

На основе предобуславливателя с LU -разложением разработана модификация, позволяющая сократить заполнение матрицы, без существенных потерь точности решения, что подтверждается экспериментально, особенно при использовании его в методе квазиминимальных невязок и методе бисопряженных градиентов. К тому же разработанный предобуславливатель с $LU(t)$ -разложением с выбором ведущего элемента применим к матрицам с произвольным портретом, СЛАО с которыми в большинстве случаев без использования предобуславливателя решить невозможно.

Литература

1. Баландин. М. Ю., Э. П. Шурина. Методы решения СЛАО большой размерности. Новосибирск: НГТУ, 2000. - 70 с.
2. Саух. С. Е. Неполная столбцово-строчная факторизация матриц для итерационного решения больших систем уравнений // Электронное моделирование / ИПМЭ им. Г.Е. Пухова НАН Украины. 2010. Т. 32. № 6. С. 3-14.
3. Harwell-BoeingCollection. [Electronic resource] // MatrixMarket: [сайт]. [1998]. URL: <http://math.nist.gov/MatrixMarket/data/Harwell-Boeing/> (5.12.17).
4. Chen T.-Y. Preconditioning Sparse Matrices for Computing Eigenvalues and Solving Linear Systems of Equations. — University of California at Berkeley, 2011. — 135 p. <http://dx.doi.org/10.1021/ja00001a054>

UDC 519.612

On an Iterative Method for Solving Ill-Conditioned Systems of Linear Algebraic Equations with Sparse Matrices

S. Y. Gogoleva, L. G. Zainetdinova

*Samara University
Moskovskoe Shosse 34 A, Samara, 443086, Russian Federation*

Email: gogoleva_s@mail.ru, zainetdinova@mail.ru

A new approach for solving systems of linear algebraic equations by iterative methods is proposed. The approach makes it possible to effectively to find an acceptable solution with a minimum filling of sparse matrices.

Key words and phrases: preconditioner, iterative methods, sparse matrices, filling, systems of equations.

УДК 519.25

Разработка модели биоимпедансометра

Д. А. Грачев, Е. Ю. Емельянова, Е. В. Попова, А. А. Хохлов

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: d.a.grachev@prezens.ru, eeu057@gmail.com, evlap.0852@gmail.com, khokhlov_aa@rudn.university

В работе рассматривается модель биоимпедансометра, генерирующего сигнал на частоте от 20 Гц до 20 МГц. Устройство состоит из двух блоков — первый блок отвечает за генерацию сигнала, второй блок отвечает за считывание и обработку полученных данных.

Ключевые слова: ВИА, биоимпеданс.

1. Введение

В современном мире большое внимание уделяется здоровью, здоровому образу жизни, состоянию тела с точки зрения физиологии. Измерение сопротивления тела помогает нам изучить его состав на тканево-клеточном уровне. Существует множество методов его измерения. В данной работе будет рассмотрен самый распространенный и доступный метод — биоимпеданс [1].

Метод реализуется через измерение сопротивления тканей организма специальным прибором — биоимпедансным анализатором.

Задача заключается в том, что бы построить прибор для проведения биоимпедансных испытаний. Устройство должно генерировать переменный ток в диапазоне от 20 Гц до 20 МГц, считывать показания тока и напряжения, по вольт-амперной характеристике производить биоимпедансный анализ. Таким образом, один блок устройства будет отвечать за генерацию переменного тока, другой — за считывание и обработку полученных данных.

2. Генератор

Генератор выполнен на основе кварцевого резонатора по схеме Пирса. Использование этой схемы позволяет получить стабильный синусоидальный сигнал. Номинальная частота кварца — 20 МГц.

Для того что бы изменять частоту в указанном диапазоне потребуются делитель на счетчиках импульсов. Для данной задачи были использованы три десятичных счетчика К155ИЕ2. Первый отвечает за единицы, второй за десятки, третий за сотни. Таким образом, мы получим делитель частоты с минимальным коэффициентом деления в 1 и максимальным в 999.

3. Обработчик

Для снятия и обработки показаний будет использоваться программируемый микроконтроллер arduino с цифровым датчиком тока и напряжения на микросхеме INA219 (рис. 1).

INA219 рассчитывает ток протекающий через шунтирующий резистор, измеряя его напряжение, а затем применяет закон Ома.

Напряжение шины измеряется непосредственно с контакта VIN-. Мощность вычисляется путем умножения этих величин.

Получение и обработка данных осуществляется на arduino с помощью сторонних библиотек.

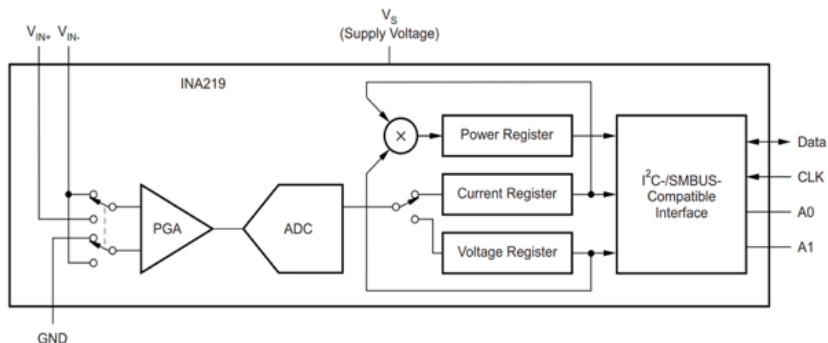


Рис. 1. Микросхема INA219

4. Заключение

В работе описано создание прибора для проведения биоимпедансных испытаний и рассмотрены его блоки. Данная модель была реализована авторами и используется для сбора данных при разработки новых математических моделей биоимпедансных исследований.

Литература

1. Мартиросов Э. Г., Николаев Д. В., Руднев С. Г., Технологии и методы определения состава тела человека. — М.: Наука, 2006.

UDC 519.25

Model of BIA Measurement Hardware

D. A. Grachev, E. U. Emelyanova, E. V. Popova, A. A. Khokhlov

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: d.a.grachev@prezens.ru, eeu057@gmail.com, evlap.0852@gmail.com, khokhlov_aa@rudn.university

The model of a biomedansometer that generates a signal at a frequency of 20 Hz to 20 MHz is considered. The device consists of two units - the first unit is responsible for the generation of the signal, the second unit is responsible for reading and processing the data.

Key words and phrases: BIA, bioimpedance.

УДК 519.25

Обработка данных акселерометров мобильных устройств для классификации движений

Гбавле Люк-Дональд Дилие, Е. Ю. Емельянова,
Е. В. Попова, А. А. Хохлов

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: lucdalie@mail.ru, eeu057@gmail.com, evlap.0852@gmail.com, khokhlov_aa@rudn.university

В работе рассматривается применение метода SSA (Singular Spectrum Analysis) к анализу показаний датчика акселерометра мобильных устройств. В результате анализа делается вывод о характере движения мобильного устройства.

Ключевые слова: SSA, акселерометр, SVD.

1. Введение

Современные мобильные устройства, такие, как смартфоны, планшеты, фитнес-браслеты и другие, практически всегда содержат в себе акселерометр. Акселерометр позволяет в каждый момент времени получать значения ускорений по каждой из осей декартовой системы координат, связанной определенным образом с корпусом устройства. Фиксируя ускорения и впоследствии анализируя их, можно делать вывод о характере движения мобильного устройства и, как следствие, человека, который это устройство носит.

Например, таким образом мобильные приложения, установленные в современных смартфонах, автоматически определяют количество шагов, которые сделал пользователь, автоматически определяют момент, когда пользователь ускоряет или замедляет шаг или переходит на бег, впоследствии составляя персональную статистику физической активности.

Более точные измерения позволяют получать специальные браслеты, которые одеваются на руку. Точность достигается за счет того, что браслет неподвижно закреплен на теле человека в отличие от смартфона, который может находиться в руке, в сумке, в кармане или где-то еще. В работе авторами рассматривается методика анализа полученных с носимых устройств данных для решения ряда прикладных задач.

2. Описание задачи

Помимо подсчета шагов и сбора статистики по физической активности, данные, полученные с акселерометра, можно использовать для решения других задач, связанных с анализом характеристик движения.

Одной из таких задач является определение момента, когда человек, в руке или в кармане которого находится смартфон, неожиданно падает. Мобильное приложение, работающее в фоновом режиме и выполняющее такую функцию, может быть установлено, например, пожилым людям или людям с заболеваниями сердца. Если человеку станет плохо и он упадет, смартфон может в автоматическом режиме отправить уведомление или позвонить родственникам. Задача анализа в данном случае – отличить падение смартфона на поверхность (например, если его бросили на диван, на стол, либо выронили на пол) от падения человека, в кармане которого находится смартфон.

Более интересной задачей является анализ качества выполнения различных спортивных упражнений. Проанализировав показания акселерометров, закрепленных на кистях и лодыжках, можно, например, определить, насколько точно человек

выполняет те или иные упражнения. Это актуально для восточных единоборств, художественной гимнастики, танцев и других видов спорта, требующих точного выполнения тех или иных движений. Авторами предлагается следующая методика анализа:

1. Показания акселерометра считываются со смартфона при помощи мобильного приложения и записываются в локальную базу данных. Периодичность получения информации может регулироваться (например, запись делается 10 раз в секунду).
2. С определенной периодичностью (например, один раз в 30 секунд) из базы данных делается выборка данных.
3. Из выборки строится трехмерных временной ряд (так как имеются данные по ускорениям по трем осям).
4. Временной ряд анализируется методом SSA [1] (при помощи троекратного применения одномерного алгоритма SSA, либо при помощи применения MSSA).
5. На основании анализа тем или иным образом делается вывод о характере движения, либо его качестве.

3. Пример решения задачи определения типа движения

Ниже приведен пример применения описанной выше методики для того, чтобы определить факт падения человека со смартфоном.

Для решения задачи были проведены эксперименты по сбору данных с акселерометра смартфона Samsung Galaxy S3, который находился в нагрудном кармане у экспериментатора. Экспериментатор многократно имитировал падение тем или иным способом, а затем, достав смартфон из кармана, экспериментатор многократно бросал его на диван, на мягкий пол и на стол. Все это время запущенное мобильное приложение собирало в фоновом режиме данные и записывало в локальную базу данных SQLite.

Впоследствии данные были проанализированы троекратным применением метода SSA для набора данных по каждому из ускорений. Ниже представлены графики распределения собственных значений в случаях падения человека со смартфоном в кармане и смартфона без человека.

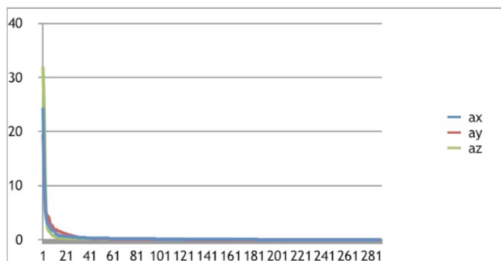


Рис. 1. Распределение собственных значений в случае падения смартфона без владельца

По вертикальной оси отложены абсолютные значения собственных значений матрицы, полученной в результате преобразования траекторной матрицы SSA. По горизонтальной оси отложены относительные периоды времени.

Хорошо видно, что значения собственных чисел отличаются для разных случаев. Это позволит в будущем создать программное обеспечение, которое с большой точностью будет определять факт падения владельца смартфона.

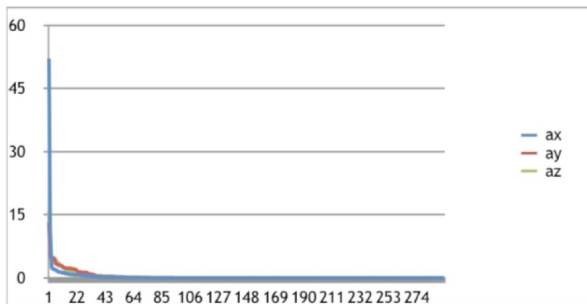


Рис. 2. Распределение собственных значений в случае падения смартфона в кармане владельца

4. Заключение

В работе авторами предлагается новая методика анализа данных с акселерометров носимых мобильных устройств для решения прикладных задач, связанных с необходимостью классификации типа или качества движения. Приведен пример решения задачи определения падения человека со смартфоном.

Литература

1. Данилов Д. Л., Жиглявский А. А. Главные компоненты временных рядов: метод «Гусеница». — Санкт-Петербургский Государственный Университет, 1997.

UDC 519.25

Accelerometer Data Analysis for User Activity Classification

Gbavle Luk-Donald Dilie, E. U. Emelyanova, E. V. Popova,
A. A. Khokhlov

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: lucdalie@mail.ru, eeu057@gmail.com, evlap.0852@gmail.com, khokhlov_aa@rudn.university

Authors proposed to apply SSA (Singular Spectrum Analysis) method of time series analysis to mobile devices accelerometer data in order to classify user activity.

Key words and phrases: computer science, information technologies, conference proceedings.

УДК 519.21;51-76

Синтез и анализ многомерных математических моделей популяционной динамики

А. В. Демидова*, О. В. Дружинина[†], О. Н. Масина[‡],
Н. Мияйлович[§], М. Ячимович[§]

* Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198

[†] Федеральный исследовательский центр «Информатика и управление» РАН
ул. Вавилова, д. 44, к. 2, Москва, Россия, 119333

[‡] Кафедра математического моделирования и компьютерных технологий
Елецкий государственный университет им. И.А. Бунина
ул. Коммунаров, д. 28, г. Елец, Россия, 399770

[§] Кафедра математики, Университет Черногории
б-р Джорджа Вашингтона, д. 2, Подгорица, Черногория, 81000

Email: demidova_av@rudn.university, ovdruzhi@mail.ru, olga121@inbox.ru, nevenamijajlovic@hotmail.com, milojica@jacimovic.me

Описано построение многомерных моделей динамики популяций, учитывающих отношения конкуренции и мутуализма. Рассмотрены модельные примеры в трехмерном и четырехмерном случаях, проведено качественное и численное исследование детерминированных моделей. Выполнен переход к соответствующим многомерным недетерминированным моделям, задаваемым с помощью дифференциальных включений, нечетких и стохастических дифференциальных уравнений, и проведен анализ устойчивости. Описана структура многомерных стохастических моделей с конкуренцией и мутуализмом. Проведен сравнительный анализ детерминированных и стохастических моделей.

Ключевые слова: многомерные модели динамики популяций, дифференциальные уравнения, одношаговые процессы, стохастическая модель, принцип редукции, конкуренция, мутуализм, устойчивость.

1. Введение

Проблема построения многомерных математических моделей популяционной динамики является актуальной в задачах исследования многовидовых сообществ [1–3]. В настоящей работе рассмотрены детерминированные и стохастические многомерные модели с учетом конкуренции и мутуализма. Изучаемые модели являются обобщениями моделей, рассмотренных в [4–8]. В [8] изучены многомерные детерминированные и стохастические модели взаимодействия популяций без учета мутуализма. В [6–8] использован метод построения самосогласованных стохастических моделей, разработанный в [9, 10], а также принцип редукции задачи об устойчивости решений дифференциального включения к задаче об устойчивости других типов уравнений.

2. Построение многомерных детерминированных моделей

В трехмерном случае рассмотрена динамическая популяционная модель «конкурент – конкурент – мутуалист»

$$\begin{aligned} \frac{dx_1}{dt} &= ax_1 \left(1 - \frac{x_1}{K_1}\right) - \frac{a\beta x_1 x_2}{1 + mx_3}, & \frac{dx_2}{dt} &= \delta x_2 \left(1 - \frac{x_2}{K_2}\right) - \eta x_1 x_2, \\ \frac{dx_3}{dt} &= \gamma x_3 \left(1 - \frac{x_3}{L_0 + Lx_1}\right), \end{aligned} \quad (1)$$

где использованы следующие обозначения: x_i — численности популяций соответственно первого конкурента, второго конкурента и мутуалиста ($i = 1, 2, 3$), γ , L_0 , L ,

$a, K_1, K_2, \beta, m, \delta, \eta$ — положительные постоянные. Модель (1) является модификацией модели, рассмотренной в [5], и характеризуется логистическим типом роста популяций конкурентов. В отсутствие мутуализма модель (1) представляет собой классическую модель Лотки–Вольтерра конкурентного взаимодействия.

В четырехмерном случае рассмотрена динамическая популяционная модель «конкурент–мутуалист–конкурент–мутуалист»

$$\begin{aligned} \frac{dx_1}{dt} &= ax_1 \left(1 - \frac{x_1}{b}\right) - \frac{cx_1x_2}{1+dx_3}, & \frac{dx_2}{dt} &= px_2 \left(1 - \frac{x_2}{q}\right) - \frac{\rho x_1x_2}{1+sx_4}, \\ \frac{dx_3}{dt} &= rx_3 \left(1 - \frac{x_3}{l+mx_1}\right), & \frac{dx_4}{dt} &= \nu x_4 \left(1 - \frac{x_4}{u+wx_2}\right), \end{aligned} \quad (2)$$

где x_1, x_2 — численности популяций первого и второго конкурентов соответственно, x_3, x_4 — численность популяций мутуалистов для x_1, x_2 в любой момент времени t . Модель (2) является частным случаем модели, рассмотренной в [4].

В n -мерном случае рассмотрена популяционная модель, описывающая динамику k видов-конкурентов, каждый из которых связан отношением мутуализма, причем каждому виду-конкуренту соответствует свой вид-мутуалиста. Предлагаемая формализация обобщенной многомерной модели имеет вид

$$\begin{aligned} \frac{dx_i}{dt} &= a_i x_i \left(1 - \frac{x_i}{b_i}\right) - \sum_{j=1, j \neq i}^k \frac{c_{ij} x_i x_j}{1 + d_{ij} x_{i+k}}, & i &= 1, \dots, k, \\ \frac{dx_i}{dt} &= \nu_i x_i \left(1 - \frac{x_i}{u_i + w_i x_{i-k}}\right), & i &= k+1, \dots, n, \quad n = 2k, \end{aligned} \quad (3)$$

где x_i ($i = 1, \dots, k$) — численности популяций видов-конкурентов, x_i ($i = k+1, \dots, n$) — численности популяций видов-мутуалистов, $a_i, b_i, c_{ij}, \nu_i, u_i, w_i$ — положительные постоянные. Размерность изучаемой модели определяется равенством $2k = n$.

3. Анализ стохастических многомерных моделей

Численный эксперимент проводился с помощью разработанного программного комплекса для численного решения систем обыкновенных и стохастических дифференциальных уравнений методами Рунге–Кутты [10]. Библиотека написана на языке python с использованием модулей numpy и scipy. Реализованы алгоритмы генерирования траекторий винеровского процесса и многоточечных распределений, аппроксимация кратных стохастических интегралов и численные алгоритмы стохастического метода Рунге–Кутты. Для детерминированных моделей «конкурент–конкурент–мутуалист» и «конкурент–мутуалист–конкурент–мутуалист» программный комплекс демонстрирует результаты, согласующиеся с аналитическими выводами.

Для моделей (1)–(3) построены соответствующие дифференциальные включения, нечеткие и стохастические дифференциальные уравнения. С помощью принципа редукции задачи об устойчивости решений дифференциальных включений к задаче об устойчивости решений нечетких дифференциальных уравнений получены условия устойчивости [11]. С помощью принципа редукции получены условия устойчивости стохастических уравнений, соответствующих построенным многомерным моделям. Доказано, что если нулевое решение нечеткого уравнения α -устойчиво по Ляпунову при каждом $\alpha \in (0, 1]$, то нулевое решение соответствующего стохастического уравнения устойчиво по вероятности. Если нулевое решение нечеткого уравнения асимптотически α -устойчиво при любом $\alpha \in (0, 1]$, то нулевое решение соответствующего стохастического уравнения асимптотически устойчиво по вероятности. Получены условия устойчивости почти наверное и устойчивости в среднем.

Результаты проведенного исследования устойчивости недетерминированных моделей популяционной динамики использованы для сравнительного анализа со свойствами моделей, построение которых было выполнено с помощью другого метода — метода построения самосогласованных стохастических моделей [8, 9]. С помощью этого метода выполнено построение стохастических моделей, соответствующих моделям (1) и многомерному обобщению модели (2). Для исследуемых систем записаны схемы взаимодействия, которые содержат символическую запись всех возможных взаимодействий между элементами системы. Для этого использованы операторы состояния системы и оператор изменения состояния системы [9].

Показано, что уравнения в моментах для полученных уравнений в форме Ланжевена совпадают соответственно с уравнениями многомерных моделей и могут служить для исследования детерминистического поведения. Исследование стохастического члена дифференциального уравнения в форме Ланжевена позволило изучить влияние введения стохастики на поведение изучаемых систем.

4. Исследование модельных примеров

Для моделей (1), (2) найдены стационарные состояния с помощью вычислительного пакета. Получены условия существования состояний равновесия, проведена оценка модельных параметров и построены фазовые портреты.

С помощью разработанного комплекса программ для стохастических моделей взаимодействия популяций были проведены численные эксперименты с выбором тех же параметров, что и для численного анализа детерминистических моделей (1) и (2). Численный эксперимент показал значительную близость траекторий для стохастического и детерминистического случаев. Как и в детерминированном случае, траектории средних значений различных реализаций выходят на стационарный режим. Анализ показал, что введение стохастики слабо влияет на устойчивость, поэтому для изучения системы можно рассматривать ее детерминистические приближение, которое определяется матрицей $A(x)$ в уравнении Фоккера–Планка.

5. Заключение

Предложенный в работе подход к синтезу моделей динамики популяций с учетом конкуренции и мутуализма позволил применить метод построения стохастических самосогласованных моделей и принцип редукции к многомерным моделям специального вида. На основе применения метода построения стохастических самосогласованных моделей осуществлен синтез соответствующих стохастических моделей. Численные эксперименты, проведенные в рамках модельных примеров, позволили провести оценку согласованности расчетов для детерминированного и стохастического случаев.

Благодарности

Работа частично поддержана грантом РФФИ № 16-01-20379. Также публикация подготовлена при поддержке программы РУДН «5-100».

Литература

1. Александров А. Ю., Платонов А. В., Старков В. Н., Степенко Н. А. Математическое моделирование и исследование устойчивости биологических сообществ. — СПб.: Соло, 2006.
2. Базыкин А. Д. Нелинейная динамика взаимодействующих популяций. — Москва–Ижевск: Ин-т компьютерных исследований, 2003.
3. Свирижев Ю. М., Логофет Д. О. Устойчивость биологических сообществ. — М.: Наука, 1978.
4. Freedman H. I., Rai B. Can mutualism alter competitive outcome: a mathematical analysis // Rocky Mountain. — 1995. — Vol. 25, no. 1. — Pp. 217–230.

5. Rai B., Freedman H. I., Addicott J. F. Analysis of three species models of mutualism in predator-prey and competitive systems // *Math. Biosci.* — 1983. — No. 63. — Pp. 13–50.
6. Demidova A. V., Druzhinina O. V., Jacimovic M., Masina O. N. Construction and analysis of nondeterministic models of population dynamics // *Communications in Computer and Information Science (CCIS)*. Springer, 2017. — P. 498–510.
7. Демидова А. В., Дружинина О. В., Масина О. Н. Анализ качественных свойств стохастических моделей популяционной динамики с конкуренцией и мутуализмом // *Материалы Пятой Национальной научной конференции с международным участием «Математическое моделирование в экологии» (ЭкоМатМод-2017), (Пушино, 16–20 октября 2017 г.)*. Пушино: ИФХ и БПП РАН, 2017. — С. 68–72.
8. Demidova A. V., Druzhinina O. V., Masina O. N. Design and stability analysis of nondeterministic multidimensional populations dynamics models // *Вестник РУДН. Серия: Математика. Информатика. Физика.* — 2017. — Vol. 25, no. 4. — Pp. 363–372.
9. Геворкян М. Н., Демидова А. В., Егоров А. Д., Кулябов Д. С., Королькова А. В., Севастьянов Л. А. Влияние стохастизации на одношаговые модели // *Вестник РУДН. Серия: Математика. Информатика. Физика.* — 2014. №1. — С. 71–85.
10. Еферица Е. Г., Королькова А. В., Геворкян М. Н., Кулябов Д. С., Севастьянов Л. А. Программный комплекс стохастического моделирования одношаговых процессов // *Вестник РУДН. Серия: Математика. Информатика. Физика.* — 2014. №3. — С. 46–59.
11. Дружинина О. В., Масина О. Н. Методы исследования устойчивости и управляемости нечетких и стохастических динамических систем. — М.: ВЦ РАН, 2009.

UDC 519.21;51-76

Synthesis and Analysis of Multidimensional Mathematical Models of Population Dynamics

A. V. Demidova*, O. V. Druzhinina[†], O. N. Masina[‡], N. Mijajlovic[§],
M. Jacimovic[§]

** Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

*† Federal Research Center «Computer Science and Control» of RAS
44-2 Vavilov str., Moscow, 119333, Russian Federation*

*‡ Department of Mathematical Modeling and Computer Technologies
Bunin Yelets State University
Communards str. 28, Yelets, Russia, 399770*

*§ Department of Mathematics, University of Montenegro
2 George Washington Blvd, Podgorica, Montenegro, 81000*

Email: demidova_av@rudn.university, ovdruz@mail.ru, olgal21@inbox.ru, nevenamijajlovic@hotmail.com, milojica@jacimovic.me

The designing of multidimensional models of population dynamics taking into account the relations of competition and mutualism is described. The model examples in three-dimensional and four-dimensional cases are considered, qualitative and numerical investigation of deterministic models is carried out. The transition to the corresponding multidimensional nondeterministic models defined by differential inclusions, fuzzy and stochastic differential equations is made, and stability analysis is performed. The structure of multidimensional stochastic models with competition and mutualism is described. The comparative analysis of deterministic and stochastic models is carried out.

Key words and phrases: multidimensional models of population dynamics, differential equations, one-step processes, stochastic model, principle of reduction, competition, mutualism, stability.

УДК 004.338

Оптимизация стоимости транзакций в блокчейне Ethereum

А. П. Денисенко*, Э .Ф. Мингажитдинова*, В. А. Петров†,
А. С. Савин*, А. А. Хохлов*

* *Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

† *Кафедра информационных технологий
Российский университет дружбы народов,
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: ggintou@gmail.com, elza.96.ru@mail.ru, vladimirap@yandex.ru, assavin90@gmail.com, khokhlov_aa@rudn.university

В работе рассматривается задача оптимизации расходов газа (*gas*) при выполнении большого количества транзакций в блокчейне Ethereum. В качестве параметра выступает максимальное время, необходимое для совершения всех транзакций.

Ключевые слова: блокчейн, ethereum, eth.

1. Введение

Блокчейн Ethereum представляет широкие возможности для разработчиков и для бизнеса [1]. В отличие от блокчейна Bitcoin, основной задачей которого является обслуживание транзакций токенов BTC между кошельками пользователей, Ethereum позволяет создавать собственные токены и программный код со сложной логикой (смарт-контракты). Возможность выпускать собственную криптовалюту (токены) легла в основу такого явления, как ICO (Initial Coin Offering). Суть ICO заключается в том, что стартапы выпускают свои токены и назначают им цену. Другие люди выступают в качестве инвесторов и покупают эти токены за другую криптовалюту либо за фиатные деньги, инвестируя таким образом в проект. Также часть выпущенных токенов стартапы раздают за различные действия пользователей, направленные на повышение известности проекта в сети (репосты, ретвиты, видеоролики и другие действия).

Обычно непосредственно смарт-контракт токенов стартапы выпускают не с самого начала ICO, а ближе к его завершению, а до этого момента количество токенов, которые впоследствии должны получить те или иные люди, фиксируется в базе данных или другим образом. В итоге, после того, как стартап выпускает токены (осуществляет загрузку в блокчейн Ethereum смарт-контракта ERC-20), встает задача рассылки нужного количества токенов инвесторам и другим пользователям.

2. Постановка задачи

В награду за майнинг блоков и выполнение на своих вычислительных мощностях кода смарт-контрактов пользователей блокчейна Ethereum, майнеры получают токен *ETH*. У этого токена есть производные единицы, минимальной из которых является *wei*. $wei = 10^{-18} ETH$. Это базовая единица для расчетов.

Если пользователь хочет прочитать какую-либо информацию из блокчейна, то такая операция ничего не стоит. Если же пользователь хочет добавить данные в блокчейн (например, записать данные о транзакции с одного кошелька на другой), то такое действие имеет определенную стоимость. Единицей стоимости транзакций в блокчейне Ethereum является газ (*gas*). Когда пользователь отправляет транзакцию в блокчейн, он указывает, сколько он максимально готов потратить газа (*GASLIMIT*) и стоимость единицы газа в *wei* (*GASPRICE*). Есть

понятие минимального количества *gas*, необходимого для осуществления транзакции, эта цифра рассчитывается в зависимости от конкретной транзакции. Таким образом, максимальная стоимость транзакции для пользователя $TRXCOST = GASPRICE * GASLIMIT$.

Например, если пользователь указал, что готов заплатить 30000 единиц *gas* ($GASLIMIT = 30000$) и стоимость одной единицы газа составляет 2000000000 *wei* ($GASPRICE = 2Gwei$), то максимальная стоимость транзакции составит $3 * 10^4 * 2 * 10^9 = 6 * 10^{13} wei$, что составляет 0,00006 *ETH*. Слово "максимальная" употребляется в том смысле, что пользователь указывает максимальное количество *gas*, которое готов потратить. Если в процессе выполнения транзакции *gas* будет исчерпан и его окажется недостаточно для выполнения транзакции, то никакие данные в блокчейн добавлены не будут, а пользователь столкнется с ошибкой. Такая ситуация встречается довольно редко, так как программное обеспечение для работы с блокчейном обычно заранее вычисляет необходимое количество *gas*.

Данные и количестве *gas* и его стоимости включаются в тело транзакции и, таким образом, программное обеспечение майнеров блокчейна может их получить. Майнеры могут настраивать свои узлы таким образом, чтобы выполнять транзакции только с определенной минимальной стоимостью газа, что зачастую и происходит. Поэтому есть прямая зависимость - чем большую стоимость *gas* готов платить пользователь, которому требуется выполнить транзакцию, тем большее количество майнеров захотят включить его транзакцию в очередной блок и, как следствие, тем быстрее выполнится транзакция. Таким образом, пользователь становится перед выбором — либо заплатить меньше и ждать выполнения транзакции определенное время, либо заплатить больше и выполнить транзакцию быстро.

Следует отметить, что существуют определенные сервисы, которые приводят в реальном времени статистику по транзакциям Ethereum, например, ethgasstation.info. Ресурс также обладает API, которое позволяет получать информацию в реальном времени. Важный параметр, который можно получить - это распределение времени ожидания выполнения транзакции в зависимости от выбранной стоимости единицы *gas*.

Итак, задача, которой посвящена данная работа, заключается в следующем - как осуществить за заданное время большое количество транзакций и при этом минимизировать их суммарную стоимость. Эта задача актуальна для различных блокчейн проектов, которые рассылают в рамках ICO или других активностей свои токены другим пользователям.

3. Алгоритм решения задачи

Для того, чтобы рассылать токены в автоматическом режиме, требуется программное обеспечение. Авторами использовалась широко известная библиотека [web3j](https://web3.js.org/), позволяющая на языке программирования Java реализовывать логику формирования, подписывания и отправки транзакций. Эта библиотека содержит функцию [web3j.ethEstimateGas](https://web3.js.org/docs/en/ethEstimateGas), которая для конкретной транзакции позволяет вычислить необходимое количество *gas* для осуществления транзакции. Также при помощи API сервисов, аналогичных ethgasstation.info, мы можем получить распределение времени выполнения транзакции в зависимости от указанной пользователем стоимости единицы *gas*. Зная эти два параметра и зная общее количество транзакций (считаем, что транзакции по переводу токенов однотипны и количество *gas* для них одинаково), можно решить задачу выбора такой стоимости *gas*, при которой все транзакции будут выполнены за заданное время (например, за 24 часа).

Однако, в данном виде решение задачи представляется некорректным, так как, во-первых, в зависимости от загруженности блокчейна, распределение времени выполнения транзакций по стоимости единицы *gas* различается, а, во-вторых, получаемая статистика не гарантирует, что, если пользователь выставит определенную стоимость *gas*, то его транзакция точно будет выполнена за указанное время.

В связи с этим, алгоритм решения задачи немного усложняется. Проверка и перерасчет стоимости единицы *gas* предлагается выполнять периодически, например каждый раз после выполнения очередного 1% от общего количества транзакций. Таким образом вероятность того, что все транзакции будут выполнены за заданное время, значительно выше, так как алгоритм будет реагировать на изменения загрузки блокчейна и измененное, как следствие этого, распределение среднего времени выполнения транзакций в зависимости от стоимости *gas*.

Данный алгоритм был применен в рамках модельных проектов в тестовой сети Ethereum Ropsten, подтвердив свою эффективность.

4. Заключение

В работе рассмотрена актуальная задача снижения стоимости рассылок токенов в блокчейне Ethereum. Как показала практика, возможно достичь экономии в 400-500% при незначительном увеличении времени рассылки.

Литература

1. Andreas M. Antonopoulos, Gavin Wood. Mastering Ethereum: Building Smart Contracts and Dapps, 1st Edition. O'Reilly Media; 1 edition (December 25, 2018).

UDC 004.338

Optimizing Transaction Cost in Ethereum Blockchain

A. P. Denisenko*, E. F. Mingazhitdinova*, V. A. Petrov[†], A. S. Savin*,
A. A. Khokhlov*

* *Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

[†] *Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: ggintou@gmail.com, elza.96.ru@mail.ru, vladimirap@yandex.ru, assavin90@gmail.com,
khokhlov_aa@rudn.university

The paper considers the problem of optimizing the cost of transactions in the Ethereum blockchain. Transaction time is limited.

Key words and phrases: blockchain, Ethereum, eth.

УДК 519.633.2

Постановка задачи расчета направляемых мод градиентного в поперечном направлении волновода

Д. В. Диваков, А. Л. Севастьянов, М. П. Белов

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: divakov_dv@rudn.university, sevastianov_al@rudn.university, mikhail.belov.pfur@gmail.com

В работе приводится постановка задачи численного расчета направляемых волноводных мод градиентных в поперечном направлении трехслойных планарных регулярных волноводов. Описан подход к построению приближенного метода решения задачи.

Ключевые слова: градиентный волновод, направляемые моды, метод Галеркина, нерегулярные краевые условия.

1. Постановка задачи

Распространение электромагнитного излучения в волноводных структурах описывается уравнениями Максвелла, материальными уравнениями и граничными условиями, выделяющими решения, соответствующие направляемым модам [1]. В планарных структурах для монохроматического излучения уравнения Максвелла можно редуцировать к двум подсистемам: для ТЕ-мод и для ТМ-мод [2]. Подсистема для ТЕ-моды в свою очередь может быть представлена в виде одного уравнения Гельмгольца относительно ведущей компоненты E_y и двух дополнительных соотношений, выражающих компоненты H_x и H_z через ведущую.

Рассматривается градиентный в поперечном направлении волновод, показатель преломления которого $n(x)$ распределен следующим образом: $n(x) = n_f(x)$, при $x \in [0; h]$; $n(x) = n_c$, при $x > h$; $n(x) = n_s$, при $x < 0$. Используя метод разделения переменных, представим искомую величину в виде $E_y(x, z) = \psi(x)V(z)$ и получим задачу на собственные значения и собственные функции оператора второго порядка с кусочно-постоянным потенциалом, пропорциональным показателю преломления волновода. Одним из подходов к решению задачи на собственные значения и собственные функции для операторов второго порядка на всей оси с потенциалом, обращаемым в константы за пределами некоторого конечного интервала, является так называемый метод волнового сопряжения. Направляемые волноводные моды удовлетворяют асимптотическому условию $E_y \rightarrow 0$ при $x \rightarrow \pm\infty$, на границах раздела сред $x = 0, h$ выполняются условия сопряжения $[E_y]_{x=0,h} = [\partial E_y / \partial x]_{x=0,h} = 0$.

Обобщая вышесказанное, задача на собственные значения и собственные функции, отвечающие направляемым волноводным модам рассматриваемого градиентного волновода, формулируется следующим образом:

$$\begin{cases} \psi'' + k_0^2 n^2(x) \psi = k_0^2 \beta^2, \\ [\psi]_{x=0,h} = [\psi']_{x=0,h} = 0, \quad \psi \xrightarrow{x \rightarrow \pm\infty} 0, \end{cases} \quad (1)$$

где k_0 — волновое число в вакууме, β — искомое собственное значение, определяющее коэффициент фазового замедления направляемой волноводной моды, ψ — искомая собственная функция, определяющая поперечную часть направляемой волноводной моды. Принимая во внимание тот факт, что при $x > h$ и $x < 0$ показатель преломления постояен можно записать общее решение уравнения из задачи (1).

Кроме того, используя асимптотические условия $\psi \xrightarrow{x \rightarrow \pm\infty} 0$, выделяющие направляемые моды, получим следующий вид решения уравнения в области покровного слоя

$$\psi|_{x>h} = A_c e^{-k_0 \sqrt{\beta^2 - n_c^2} (x-h)}$$

и аналогичный вид решения в области подложки

$$\psi|_{x<0} = A_s e^{k_0 \sqrt{\beta^2 - n_s^2} x}.$$

Вид решения в области, соответствующей волноводному слою, в случае произвольной зависимости $n_f(x)$ не известен. Используя вид решения, соответствующего направляемым модам при $x > h$ и $x < 0$ и условия сопряжения $[\psi]_{x=0,h} = [\psi']_{x=0,h} = 0$ можно сформулировать условия третьего рода на границах волноводного слоя, тем самым редуцировав задачу на собственные значения и собственные функции с асимптотическими условиями Дирихле на всей оси (1) к следующей задаче на отрезке $x \in [0, h]$ с условиями третьего рода с нелинейным вхождением спектрального параметра с граничные условия:

$$\begin{cases} \psi'' + k_0^2 n_f^2(x) \psi = k_0^2 \beta^2, \\ \left(\psi' - k_0 \sqrt{\beta^2 - n_s^2} \psi \right)_{x=0} = 0, \quad \left(\psi' + k_0 \sqrt{\beta^2 - n_c^2} \psi \right)_{x=h} = 0. \end{cases} \quad (2)$$

2. Построение численного решения

В силу произвольности потенциала $n_f(x)$ в задаче (2) решение необходимо строить численно. Краевые задачи для уравнений второго порядка эффективно решаются и разностными методами, и проекционными методами Галеркина. В случае, когда задача на собственные значения и собственные функции является вспомогательной и искомые собственные функции предполагается использовать как базис для разложения по нему приближенного решения более сложной задачи для уравнения в частных производных удобнее применять метод Галеркина [3]. В таком случае, получаемые собственные функции можно интегрировать и дифференцировать в символьном виде, что позволяет строить символьно-численные методы расчета, которые в исследованиях авторов показали и вычислительную эффективность, и большую физическую наглядность получаемого решения.

Трудность построения численного метода решения задачи (2) методом Галеркина состоит в нелинейных граничных условиях, в которые входит собственное значение β . Решение задачи (2) методом Галеркина предполагает разложение искомой величины ψ по полной системе функций, удовлетворяющих граничным условиям из (2). В случае краевых условий Дирихле и Неймана известны полные системы функций. Построение полной системы для задач с краевыми условиями третьего рода без нелинейного вхождения в них собственного значения также возможно с применением численных методов. Однако построить систему функций для разложения по ней величины ψ — задача не менее сложная, чем рассматриваемая задача (2). Отметим также, что нелинейные условия третьего рода характерны для волноводных задач, например, похожие условия встречаются в задачах отыскания вытекающих мод [2]. Предлагается заменить искомую функцию ψ на вспомогательную функцию φ , такую чтобы в новой задаче относительно φ коэффициенты в граничных условиях не зависели от β , кроме того, желательно перейти от условий третьего рода к более простым — Дирихле или Неймана. Введем φ следующим образом:

$$\psi(x) = e^{px^2/2+qx} \phi(x), \quad (3)$$

где p и q будем подбирать таким образом, чтобы избавиться от нелинейных коэффициентов в граничных условиях. С учетом (3) граничные условия для вспомогательной функции φ имеют следующий вид:

$$\left(q - k_0 \sqrt{\beta^2 - n_s^2}\right) \varphi(0) + \varphi'(0) = 0, \quad \left(ph + q + k_0 \sqrt{\beta^2 - n_c^2}\right) \varphi(h) + \varphi'(h) = 0. \quad (4)$$

Параметры p и q подбираем таким образом, чтобы коэффициенты в граничных условиях (4) не зависели от β . Очевидным выбором являются следующие значения $q = k_0 \sqrt{\beta^2 - n_s^2}$, $p = -\frac{1}{h} \left(k_0 \sqrt{\beta^2 - n_s^2} + k_0 \sqrt{\beta^2 - n_c^2}\right)$, которые позволяют сформулировать для φ задачу с однородными условиями Неймана $\varphi(0) = \varphi(h) = 0$. Преобразование (3) усложнит дифференциальный оператор в задаче (2), однако, благодаря регулярным граничным условиям построить приближенное решение новой задачи относительно φ не составляет труда.

3. Заключение

В работе рассмотрена одна из возможных постановок задачи на собственные значения и собственные функции, определяющие направляемые моды градиентного волновода. Описан возможный подход к решению задачи с нерегулярными краевыми условиями, зависящими от спектрального параметра. Предложена замена задачи на вспомогательную с регулярными краевыми условиями, для которой построение приближенного решения методом Галеркина осуществляется очевидным образом.

Литература

1. Могилевский И. Е., Свешников А. Г. Математические проблемы теории дифракции. М.: МГУ, 2010.
2. Самарский А. А., Тихонов А. Н. О представлении поля в волноводе в виде суммы полей ТЕ и ТМ // ЖТФ. — 1948. — Т. 18. — Вып. 7. — С. 959–970.
3. Диваков Д. В., Малых М. Д., Севастьянов А. Л., Севастьянов Л. А. Моделирование распространения поляризованного света в тонкопленочной волноводной линзе // Вестник РУДН. Серия: Математика. Информатика. Физика. — 2017. — Т. 25, № 1. — С. 56–68.

UDC 519.633.2

Statement of the Problem of Calculation of the Guided Modes of the Gradient in the Transverse Direction Waveguide

D. V. Divakov, A. L. Sevastianov, M. P. Belov

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: divakov_dv@rudn.university, sevastianov_al@rudn.university, mikhail.belov.pfur@gmail.com

The paper presents the formulation of the problem of numerical calculation of the guided waveguide modes of gradient transverse directions of three-layer planar regular waveguides. An approach to the construction of an approximate method for solving a problem is described.

Key words and phrases: gradient waveguide, guided modes, the Galerkin method, irregular boundary conditions.

УДК 519.6

Задача символьно-численного расчета собственных значений вытекающих мод в регулярном однородном открытом волноводе

А. С. Древицкий, Д. В. Диваков, М. М. Котюков

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: adrevitskiy@gmail.com, divakov_dv@rudn.university, mmkotyukov@gmail.com

В настоящей работе описан алгоритм нахождения собственных значений вытекающих мод в плоском трехслойном диэлектрическом волноводе. В работе исследуется дисперсионная зависимость первого коэффициента фазового замедления, мнимая часть которого является наименьшей, от толщины волновода. Для поиска коэффициентов фазового замедления был использован метод Хука-Дживса. Описанный алгоритм символьно-численного расчета реализован в системе компьютерной алгебры Maple.

Ключевые слова: вытекающие моды, собственные значения, дисперсионная зависимость, символьно-численный расчет.

1. Введение

В ходе настоящего исследования рассмотрена задача нахождения собственных значений для вытекающих волноводных мод в диэлектрическом волноводе.

Математической постановкой задачи для вытекающих мод, описанной выше, выступает [1, 2]:

$$\begin{cases} \psi_f'' + k_0^2 n_f^2 \psi_f = k_0^2 \beta^2 \psi_f \\ \psi_f'|_{x=h} = ik_0 \sqrt{n_c^2 - \beta^2} \psi_f|_{x=h}, \\ \psi_f'|_{x=0} = -ik_0 \sqrt{n_s^2 - \beta^2} \psi_f|_{x=0}, \end{cases} \quad (1)$$

где ψ_f — искомое решение уравнения для волноводного слоя, β — собственное значение, k_0 — волновое число в вакууме, n_c — показатель преломления покровного слоя, n_f — показатель преломления волноводного слоя, n_s — показатель преломления подложки, h — толщина волноводного слоя.

2. Основная часть

Рассматриваемым объектом данной задачи является планарный трехслойный открытый волновод с параметрами $n_c = 1$, $n_f = 2.10$, $n_s = 1.47$, $\lambda = 0.55$ нм., $h = 2\lambda$, структура которого представлена на рис. 1.

Перейдем к решению задачи (1). Решение уравнения для поперечной части должно удовлетворять условиям сопряжения. Общее решение уравнения из (1) имеет следующий вид:

$$\psi_f(x) = A_f e^{ik_0 \sqrt{n_f^2 - \beta^2} x} + B_f e^{-ik_0 \sqrt{n_f^2 - \beta^2} x}. \quad (2)$$

Подставляя данный вид решения (2) в граничные условия из (1), получаем однородную систему линейных алгебраических уравнений:

$$\begin{cases} iA_f p_f e^{ip_f h} - iB_f p_f e^{-ip_f h} - ip_c (A_f e^{ip_f h} + B_f e^{-ip_f h}) = 0, \\ iA_f p_f - iB_f p_f + ip_s (A_f + B_f) = 0, \end{cases} \quad (3)$$

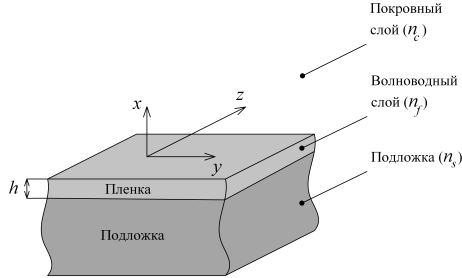


Рис. 1. Схема плоского трехслойного диэлектрического волновода

где $p_f = k_0 \sqrt{n_f^2 - \beta^2}$, $p_c = k_0 \sqrt{n_c^2 - \beta^2}$ и $p_s = k_0 \sqrt{n_s^2 - \beta^2}$.

Для того, чтобы система (3) имела нетривиальное решение необходимо, чтобы определитель матрицы коэффициентов системы был равен нулю:

$$\det \mathbf{M}(\beta) = 0. \quad (4)$$

Решениями уравнения (4) будут являться допустимые значения коэффициента фазового замедления β_j : такое комплексное число, при котором мнимая и действительная части определителя будут равняться приближенно нулю.

С помощью метода Хука–Дживиса на промежутке $0 < \text{Re}\beta_j < n_s$ и $0 < \text{Im}\beta_j < \infty$ находим корни уравнения (4): найдем минимумы функции $|\det \mathbf{M}(\beta)|^2$ [3].

График дисперсионной кривой показывает зависимость β от толщины волноводного слоя. На рис. 2 представлены графики дисперсионных кривых действительной и мнимой частей β^* , мнимая часть которой является наименьшей из всех β , для разных толщин волноводного слоя.

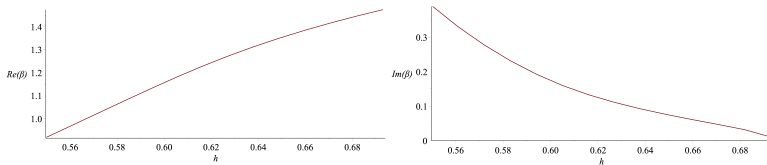


Рис. 2. Графики дисперсионных кривых β^*

Ниже (см. табл. 1) представлены значения β^* при произвольно выбранной толщине волноводного слоя.

3. Заключение

В рамках настоящей работы был представлен алгоритм нахождения собственных значений вытекающих мод. В качестве метода решения был выбран метод Хука–Дживиса. Произведен численный расчет дисперсионной зависимости собственных значений слабовытекающих мод от толщины волноводного слоя.

Таблица 1

Таблица значений β^*

h	Действительная часть β^*	Мнимая часть β^*
0.55	0.9187135321102090	0.3910511566893390
0.583	1.076580212412647	0.2318896583970420
0.627	1.271199683533986	0.1106532452107190
0.693	1.473693982820379	0.009583644457937000

Литература

1. Zhu J., Lu Y. Y. Leaky Modes of Slab Waveguides — Asymptotic Solutions // Journal of Lightwave Technology, 2006.
2. Emami F., Hoseini F. Leaky Mode Analysis of Non-Homogeneous Asymmetric Slab Waveguides // International Journal of Fundamental Physical Sciences, 2013.
3. Сотский А. Б., Steingart L. M., Jackson J. H., Чудаковский П. Я., Сотская Л. И. Призменное возбуждение вытекающих мод тонких пленок // Журнал технической физики, 2013.

UDC 519.6

The Problem of Symbolic-Numeric Computation of the Eigenvalues of the Leaky Modes in a Regular Homogeneous Open Waveguide

A. S. Drevitskiy, D. V. Divakov, M. M. Kotyukov

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: adrevitskiy@gmail.com, divakov_dv@rudn.university, mmkotyukov@gmail.com

In this paper we describe an algorithm for finding the eigenvalues of the leaky modes in a planar three-layer dielectric waveguide. The dispersion dependence of the first phase deceleration coefficient, the imaginary part of which is the smallest one, on the thickness of the waveguide is investigated in the paper. To find the phase deceleration coefficients, the Hook-Jeeves method was used. The described algorithm of symbolic-numeric computation is implemented in the Maple computer algebra system.

Key words and phrases: leaky modes, eigenvalues, dispersion dependence, symbolic-numeric computation.

УДК 004.4

Исследование свойств жидкокристаллических волноводных структур

А. А. Егоров

Институт общей физики им. А.М. Прохорова РАН

Email: yegorov@kapella.gpi.ru

Численным моделированием и экспериментально исследованы оптические волноводящие структуры из нематического жидкого кристалла.

Ключевые слова: численное моделирование, волновод, планарная линза, лазерное излучение, жидкий кристалл, директор, флуктуации, нерегулярности.

1. Введение

В настоящей работе исследованы интегрально-оптические волноводы и тонко-слоистые волноводные линзы, образованные с помощью нематического жидкого кристалла (НЖК), в том числе при наложении мощного импульсно-периодического электрического поля [1, 2]. Актуальность и перспективность этих исследований обусловлена необходимостью разработки новых методов исследования волноводящих структур из жидких кристаллов (ЖК). Кроме того, есть потребность в разработке новых методов исследования нерегулярностей жидкокристаллических волноводных слоев.

2. Экспериментальная установка. Волноводящие НЖК-структуры

Объекты исследования – интегрально-оптический волновод (рис. 1, 2), образованный из НЖК 4-Циано-4'-пентилбифенил (4-Cyano-4'-pentylbiphenyl или 5CB) хорошо известный по публикациям в научной и научно-технической литературе [1–3], а также – планарные волноводящие НЖК-линзы (рис. 1 (б)).

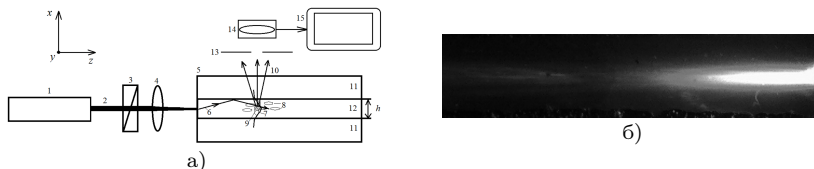


Рис. 1. а) Схематическое изображение экспериментальной установки; б) Фотография треков волноводных мод: слева и в центре фото – в интегрально-оптическом НЖК-волноводе; справа планарная НЖК-линза

На рис. 1 (а) обозначены: 1 – лазер, 2 – лазерный луч, 3 – поляризатор, 4 – линза, 5 – НЖК-волновод, 6 – распространение волноводной моды (в оптико-лучевом представлении), 7 – вертикальный профиль направляемой основной ТЕ-моды, 8 – схематическое изображение молекул ЖК, 9 – рассеивающий объем в НЖК-волноводе, 10 – рассеянное в НЖК-волноводе лазерное излучение, 11 – стеклянные пластинки, 12 – слой ЖК, 13 – щель, 14 – фотокамера, 15 – компьютер.

На рис. 1 (б) справа на самом краю фотографии видна планарная НЖК-линза, сфокусированная у торца волновода. Сразу за линзой хорошо виден характерный сфокусированный ею пучок излучения и треки мод. Размеры линзы: радиус 1,5 мм, толщина в центре определялась толщиной НЖК-волноводов (см. ниже).

3. Результаты исследования НЖК-структур

В экспериментах исследовались многомодовые НЖК-волноводы, образованные двумя стеклянными пластинами и ЖК слоем между ними (рис. 1). НЖК имел однородную планарную ориентацию с оптической осью вдоль направления директора (совпадает с осью z). В экспериментах исследовались волноводы с различными толщинами (25, 75 и 125 мкм), но основная часть результатов дана только для НЖК-волноводов с $h = 25$ и 75 мкм. Расстояние между медными электродами: 2 мм. На них подавалось высоковольтное импульсное (частота повторения 10 Гц) напряжение. НЖК слой имел показатели преломления: обыкновенный $\approx 1,53$ и необыкновенный $\approx 1,70$ (для длины волны лазерного излучения $\lambda \approx 0,64$ мкм и температуры $\approx 25^\circ\text{C}$). Стеклянные пластинки имели показатели преломления: $n_1 = n_3 \approx 1,52$. Таким образом, в экспериментах и численных расчетах исследовались трехслойные симметричные волноводы со следующими показателями преломления: $n_1 = n_3 = 1,52$, $n_2 = 1,53$. Эффективность ввода излучения в НЖК-волноводы не превышала 25–35%.

С целью количественной оценки оптических потерь в волноводе проводилось компьютерное фотометрирование фотографий соответствующих треков мод. Для определения оптических потерь использовалось известное выражение:

$$I(z) = I_0 \exp(-\alpha z), \quad (1)$$

где $I(z)$ — интенсивность излучения (на единицу площади) в любой точке по длине волновода, I_0 — интенсивность при $z = 0$.

Приведем для примера результаты компьютерной обработке данных для некоторых НЖК-волноводов. ТЕ-поляризация: без поля — $\bar{\alpha}_{\text{ТЕ}} = 4,5\text{см}^{-1}$, $r_{\text{ТЕ}} \approx 0,15 - 0,6$ мкм; с полем — $\bar{\alpha}_{\text{ТЕ}} = 3\text{см}^{-1}$, $r_{\text{ТЕ}} \approx 0,1 - 0,4$ мкм. ТМ-поляризация: без поля — $\bar{\alpha}_{\text{ТМ}} = 6\text{см}^{-1}$, $r_{\text{ТМ}} \approx 1,0 - 1,8$ мкм; с полем — $\bar{\alpha}_{\text{ТМ}} = 5\text{см}^{-1}$, $r_{\text{ТМ}} \approx 0,5 - 1,5$ мкм.

Оценка размеров НЖК-нерегулярностей проводилась в соответствии с критерием Рэлея. Определив в соответствии с формулой (1) величину потерь α в волноводе оптической мощности, можно оценить эффективные размеры r_{ef} нерегулярностей ЖК слоя в соответствии с приближенной формулой: $r_{ef} \approx \lambda \alpha h_{ef} (4\pi)^{-1} C(\theta)$, где $C(\theta)$ — безразмерный поправочный коэффициент.

Оценка для ряда НЖК-волноводов, выполненная в соответствии с критерием Рэлея, показала, что размеры нерегулярностей НЖК слоя изменяются в диапазоне от примерно 0,08 до 2,00 мкм, при этом погрешность определения не превышает 15–20%.

Известные закономерности рассеяния света в не волноводном случае объяснены на основе рассмотрения флуктуаций локальной ориентации молекул, т.е. директора НЖК [3]. В результате может быть вычислен радиус корреляций ξ флуктуаций директора, который показывает, на каком расстоянии сохраняется порядок, заданный, например, стенкой (подложкой волновода) при наложении поля E (чем больше поле, тем на меньшем расстоянии от стенки сохраняется заданный ею порядок). Он может быть поставлен в соответствие найденному радиусу корреляции нерегулярностей НЖК-волновода. Сравним предсказываемую теорией жидких кристаллов закономерность изменения радиуса корреляций флуктуаций директора $\xi \propto E^{-1}$ с полученными нами результатами для изменения радиуса корреляции нерегулярностей НЖК-волновода без поля и с полем: радиус корреляции при включении внешнего поля уменьшается для обоих типов мод. Это говорит о хорошем соответствии полученных результатов классической теории флуктуаций директора в ЖК [3]. Итак, полученная оценка подтверждает правильность обнаруженных нами экспериментальных зависимостей. Поскольку $I(z) \propto \exp(-\alpha z) \propto \exp(-r_{ef} z)$, то при уменьшении радиуса корреляции будут уменьшаться как I , так и коэффициент затухания, что и обнаружено в экспериментах.

На рис. 2 приведен один из измеренных (кривая 1) профилей интенсивности излучения недалеко от задней фокальной плоскости планарной ЖК-линзы. На

рис. 2 также обозначены: 2 — кривая, полученная сглаживанием распределения 1; 3 — подгоночная гауссова кривая; 4 — обозначает величину погрешности (менее 7%). Полуширина распределений типа 1 в окрестности фокальной плоскости, измеренная по уровню 0,5 от его максимума, доходила до $(10-45)\lambda$; это обусловлено тем, что линзы имели малую апертуру и исследовались многомодовые (14 и более мод) ЖК-волноводы.

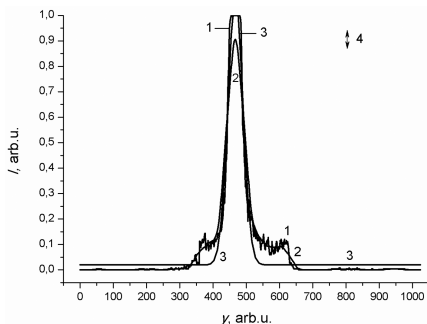


Рис. 2. Профиль интенсивности лазерного излучения за планарной НЖК-линзой

Оценка фокусного расстояния исследованных жидкокристаллических линз показала, что оно находится в диапазоне от $2R$ до $8R$, где R — радиус планарной ЖК линзы.

Литература

1. Ayriyan A. A., Ayrjan E. A., Egorov A. A. et al. Some Features of Second Harmonic Generation in the Nematic Liquid Crystal 5CB in the Pulsed-Periodic Electric Field // *Physics of Wave Phenomena*. — 2016. — Vol. 24, no. 4. — Pp. 259–267.
2. Egorov A. A. Influence of fluctuations of local orientation of nematic liquid crystal molecules on the coefficient of damping of waveguide modes // *Proc. of the VI International conference “Problems of Mathematical Physics and Mathematical Modelling”*, Moscow, MEPhI, May 25–27, 2017. — Moscow: MEPhI, 2017. — Pp. 128–129.
3. Блинов Л.М. Электро- и магнитооптика жидких кристаллов. — М.: Наука, 1978.

UDC 004.4

Research of Properties of the Liquid Crystal Waveguide Structures

A. A. Egorov

General Physics Institute, Russian Academy of Sciences

Email: yegorov@kapella.gpi.ru

Nematic liquid crystal optical waveguide structures were investigated by the numeric simulation and experimentally.

Key words and phrases: numerical simulation, waveguide, planar lens, laser radiation, liquid crystal, director, fluctuations, irregularities.

УДК 004.4

Анализ востребованных форматов заданий для онлайн-курсов среди студентов разных учебных заведений

П. А. Зотова*, Л. В. Жукова†, Д. А. Пяткина*

* Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198

† Департамент прикладной экономики
НИУ Высшая школа экономики
ул. Мясницкая, д. 20, Москва, Россия, 101000

Email: zotova-pa@mail.ru, lvzhukova@hse.ru, pyatkina_da@rudn.university

В данной работе предполагается проанализировать и выявить наиболее востребованные форматы заданий для онлайн-курсов среди студентов при использовании данных, полученных при помощи анкетирования учащихся экономического колледжа Москвы и студентов 1 курса Российского университета дружбы народов.

Ключевые слова: опрос, онлайн-курс, формат задания.

1. Введение

Многие учебные заведения сегодня вслед за ведущими мировыми разрабатывают и внедряют программу дистанционного обучения студентов с использованием онлайн-курсов. Однако прямой перенос в связи со спецификой нашего образования не возможен и требуется учитывать многие особенности. Поэтому выявление наиболее востребованных форматов заданий для онлайн-курсов актуально для многих учебных заведений и преподавателей.

Для решения поставленной задачи было проведено маркетинговое исследование в формате опроса для проверки гипотез и предположений.

Предположение: существует целевая группа, в которой востребовано дополнительное образование, и существует группа, для которой нет интереса в онлайн курсах.

2. Анализ данных

Для проведения опроса была разработана анкета из 23 вопросов, которая включала в себя три блока: информация о респонденте и его отношении к учёбе, опыт использования онлайн-курсов, наиболее интересные форматы заданий [2]. Было опрошено 100 студентов, 59 из которых являются учениками экономического колледжа и 41 студентами первого курса Российского университета дружбы народов.



Рис. 1

При анализе данных использовались методы математической статистики, такие как графики, таблицы, гистограммы.

Сначала был проведён анализ получившейся выборки респондентов. Были сделаны следующие выводы.

Выводы относительно к учебе, на основании которых формируется общий портрет современного учащегося колледжа или студента 1 курса университета [1]:

- Большая часть студентов (42%) колледжа называют главной причиной того, что они не пошли учиться в 11 класс, желание быстрее получить образование и начать работать, при этом 23% студентов отмечают, что хотели бы после окончания колледжа получить ещё и высшее образование.
- Большая часть студентов (31%) отвечают, что учат в университете или колледже то, что им интересно.
- Наиболее популярными социальными сетями и мессенджерами среди молодёжи являются «ВКонтакте» (26%), Instagram (23%), WhatsApp (22%).
- Половина всех опрошенных считают, что учатся на среднем уровне по сравнению со своими сверстниками.

Выводы относительно успеваемости:

- Гипотеза о том, что чем выше успеваемость студента, тем больше вероятность того, что в свободное время он предпочитает заниматься чем-то ещё, подтвердилась. К тому же можно заметить, что большая часть отличников (29%) в свободное от учёбы время принимают участие в каких-либо конкурсах или олимпиадах.
- Большая часть студентов, занимающихся на «отлично» (56%), с уверенностью отвечают, что знают свои цели на ближайшие 5 лет; большая часть студентов, занимающихся на уровне выше среднего (58%) и на среднем уровне (48%), отвечают, что думают об этом серьёзно, но пока не определились окончательно.
- Независимо от успеваемости, большая часть респондентов отвечают, что готовятся к контрольным работам ближайшие пару дней. Хотя среди отличников находятся те, кто предпочитают начинать готовиться к итоговой проверке знаний заранее и последовательно (44%).
- Наиболее популярным ответом про причины неиспользования онлайн-курсов является вариант ответа про недостаток мотивации. Его выбирают 44% опрошенных.

	Отличники и хорошисты	Ученики с успеваемостью на среднем уровне и ниже	Количество студентов, выбравших данный ответ
Недостаточно мотивации	40,9%	59,1%	44
Не нравится, что нужно проходить весь курс, а не выборочные темы	34,8%	65,2%	23
Не нравится, что часто ограничивается срок прохождения курса	53,3%	46,7%	15
Нет возможности так часто пользоваться Интернетом	63,6%	36,4%	11
Я использую онлайн-курсы	85,7%	14,3%	7
Всего			100

Рис. 2

Были выделены следующие наиболее популярные дисциплины для онлайн-обучения: финансы (19%), английский язык (14%), экономика (14%).

На вопрос про цели использования онлайн-курсов среди студентов были получены следующие ответы: 30% говорят, что хотели бы научиться решать определённый тип задач, 29% выбрали бы онлайн-курсы для изучения чего-то нового, что не проходят в своём учебном заведении, 21% хотели бы готовиться к экзаменам или олимпиадам и 20% для закрепления знаний, полученных в своём учебном заведении.

Далее была рассмотрена подзадача, которая включала в себя анализ группировок студентов.

Отличники говорят, что стали бы использовать онлайн-курсы в первую очередь для изучения решения определённого типа задач (55%), хорошисты выбирают ответ «для изучения новых тем, которые не раскрываются в учебном заведении» (36%), студенты, успевающие на удовлетворительную оценку, выбрали бы онлайн-курсы тоже для изучения определённого типа задач (30%).

3. Вывод

На основе данного первоначального анализа был сделан вывод, что при разработке нового онлайн-курса следует ориентироваться больше всего на студентов, успевающих в своих учебных заведениях на удовлетворительную оценку и выше. Наиболее востребованными дисциплинами для всех студентов являются экономика и финансы. Большинство студентов стали бы использовать онлайн-курсы для изучения определённого типа задач или для изучения каких-то новых тем и разделов. Продвигать онлайн-курсы было бы разумнее всего в социальной сети Вконтакте, так как её выбирают 97% опрошенных.

Литература

1. Магнус Я. Р. Эконометрика. Начальный курс: Учебное пособие для вузов / Я. Р. Магнус, П. К. Катышев, А. А. Пересецкий. — 2-е изд., испр. — М.: Дело, 1998. — 248 с.: ил.
2. Ядов В. А. Стратегия социологического исследования. — М.: Академкнига ИКЦ, 2003. — 596 с.

UDC 004.4

Analysis Vostroyans Popular Formats of Assignments for Online Courses Among Students of Different Educational Institutions

P. A. Zotova*, L. V. Zhukova[†], D. A. Pyatkina*

* *Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

[†] *Department of Applied Economics
NRU Higher School of Economics
20 Myasnitskaya str., Moscow, 101000, Russian Federation*

Email: zotova-pa@mail.ru, lvzhukova@hse.ru, pyatkina_da@rudn.university

In this paper, it is planned to analyze and identify the most popular formats of tasks for online courses among students using the data obtained by means of questionnaires of students of the Economic College of Moscow and 1st year students of RUDN University.

Key words and phrases: survey, online course, format of task.

УДК 519.864.3

Оптимальные веса рискованных активов в индексных портфелях постоянной пропорции

Ю. Ф. Касимов, М. И. Тимербаев

*Департамент анализа данных, принятия решений и финансовых технологий
Финансовый университет при Правительстве РФ
Ленинградский просп., д. 49, ГСП-3, Москва, Россия, 125993*

Email: y.f.kasimov@mail.ru, timerbayevgroup@yandex.ru

Работа посвящена анализу оптимальных индексных портфелей в стратегиях постоянной пропорции. Основным моментом в этих стратегиях является выбор пропорции в которой инвестор распределяет свой капитал между рискованными и безрисковыми активами. На основе исторических данных (ex-post) определяются оптимальные веса фондов в индексных портфелях. Рассмотрены реализации индексных портфелей для пяти наиболее крупных российских индексных фондов. Было обнаружено, что практически во всех случаях оптимальные веса фондов принимали экстремальные значения 0 или 1. При этом, частоты этих значений в выбранной последовательности примерно одинаковы. Этот эмпирический факт можно условно назвать принципом экстремальности или принципом «все или ничего».

Ключевые слова: пассивное инвестирование, стратегии постоянной пропорции, индексы, индекс ММВБ, индексные портфели, оптимальные веса.

1. Введение

Как известно в практике инвестирования используются два основных вида стратегий : активные и пассивные. Активные стратегии предполагают постоянный анализ состояния рынка, прогнозирование цен (доходностей) активов и перестройку портфеля (покупку/продажу) активов. Такие стратегии весьма дорогие (комиссии, плата за получение обработку и анализ информации и, собственно, за управление).

Пассивные (индексные) стратегии заключаются в копировании индекса [1–3]. Такие стратегии систематически реализуют так называемые индексные фонды, аккумулирующие средства инвесторов (вкладчиков фонда) и инвестирующих их в широко диверсифицированный рыночный портфель активов. Как правило, представителем рыночного портфеля является некоторый фондовый индекс. В России это, прежде всего, индексы ММВБ и, в меньшей степени, РТС. Индексный портфель копирует какой либо индекс в том смысле, что веса активов в индексном портфеле соответствуют весам этих активов в индексе. Для капитализированных индексов таких как S&P500, и ММВБ вес акции в индексе просто равен ее доли в общей рыночной капитализации [1]. Подробнее об индексах см. [4].

2. Принципы оценивания стратегий постоянной пропорции

Толчком к появлению и распространению таких стратегий было создание в 60-70 годах прошлого столетия группой американских экономистов (Тобин, Шарп, Моссин и др.) так называемой Модели Оценивая Активов или CAPM в английской аббревиатуре (Capital Asset Price Model) [5,6].

В рамках модели CAPM Тобин доказал свою знаменитую теорему о двух фондах: Любой эффективный и, значит, оптимальный по доходности или риску, портфель π есть комбинация безрискового актива с весом w_0 и рыночного портфеля с весом w_M . При этом доходность такого и риск такого портфеля определяются формулами: (см. [5]):

$$r_\pi = w_0 r_0 + w_M r_M \quad \text{и} \quad \sigma_\pi = w_M \sigma_M.$$

Теорема Тобина о двух фондах сыграла важную роль в инвестиционном бизнесе. Стало ясно, что если создать хорошо диверсифицированный фонд, размещающий капитал в широкий спектр активов, то проблема оптимального построения портфеля для любого инвестора сведется просто к выбору *оптимальной* для данного инвестора пропорции между его *безрисковой* и *рисковой* (рыночной) частью, т.е. к определению двух весов w_0 и w_M связанных очевидным соотношением: $w_0 + w_M = 1$. Такими фондами и стали как раз индексные фонды.

Но остается проблема выбора этих весов. В широко распространенных стратегиях постоянной пропорции инвестор просто выбирает тот или иной уровень вложения в рискованные активы. Выбрав долю, вкладываемую в рыночный портфель (на практике это какой либо индексный фонд), он поддерживает этот уровень *постоянным* в течение достаточно большого инвестиционного периода. Поскольку стоимость пая индексного фонда постоянно меняется, то меняется и пропорция вложенного в индексный фонд капитала. Для ее поддержания необходимо ребалансировать портфель, т.е. в случае роста доли фонда перевложить часть капитала вложенного в фонд в безрисковый актив, а в случае снижения наоборот за счет части капитала вложенного в безрисковый актив (например депозит в надежном банке) купить дополнительные паи индексного фонда. Период, в начале которого осуществляется эта операция называется периодом ребалансирования. Поскольку данные о стоимости паев индексных фондов, а следовательно и их доходности известны, то можно смоделировать *ex post* стратегии вложения в фонды для разных значений пропорций на различных инвестиционных периодах и различных периодах ребалансирования.

Целью настоящей работы и было на реальных исторических данных для российских индексных фондов (на основе индекса ММВБ) определить распределение значений оптимальных весов этих фондов для широкого спектра инвестиционных горизонтов и различных периодов ребалансирования.

Как указывалось выше целью работы являлась оценка и анализ оптимальных стратегий с индексными портфелями российских индексных фондов. Для исследования были отобраны 4 самых крупных по чистой стоимости активов, работающих на рынке не менее 10 лет и «хорошо отслеживающих» индекс ММВБ (с коэффициентом детерминации R^2 и бетой близкими к 1) индексных фондов: РГС, Открытие, ВТБ, Райффайзен, и один крупный фонд Альфа-капитал с довольно низкими коэффициентами детерминации (R^2) и бетой.

По данным о ежедневных котировках стоимости паев этих фондов за указанный период была сформирована *база данных* месячных доходностей этих фондов. Кроме того, в эту базу данных были включены данные о безрисковых ставках (годовых, полугодовых, квартальных и месячных) за этот период. Была также разработана универсальная программа, определяющая оптимальные веса индексного фонда и соответствующие коэффициенты роста портфеля постоянной пропорции для различных периодов инвестирования и подпериодов ребалансирования.

3. Основной результат исследования

В результате анализа оказалось, что *для всех* рассмотренных фондов *на всех* инвестиционных периодах оптимальные веса индексных портфелей почти во всех случаях принимают *экстремальные значения*: 0 или 1! Промежуточные значения встречаются исключительно редко. Для всех фондов и инвестиционных периодов, коэффициенты роста индексного портфеля всегда монотонно зависят от веса фонда. При этом нулевому значению оптимального веса (полное вложение в безрисковый актив) соответствует полное вложение капитала портфеля в безрисковый актив, а единичное значение — полному вложению капитала портфеля в индексный фонд. Поэтому принцип оптимальности весов индексных портфелей можно назвать (условно) принципом «все или ничего».

Конкретные значения коэффициентов роста меняются от фонда к фонду и зависят как от точности копирования индекса (близости коэффициентов R^2 и беты

к единице), комиссии взвешиваемой фондом и др. Но общая картина *экстремального распределения оптимальных весов* во всех случаях одна и та же. Совершенно аналогичной она и в случае «идеального» индексного портфеля, когда в качестве рыночной доходности берется непосредственно доходность индекса ММВБ.

Этот результат показывает, что, строго говоря, стратегия постоянной пропорции с весом вложения в индекс отличным от 0 и 1 практически не бывает оптимальной. Но это не означает практическую нецелесообразность использования стратегии постоянной пропорции. Наше исследование опиралось на использовании прошлых (исторических) данных. При планировании структуры портфеля для будущих инвестиционных периодов, заранее невозможно предсказать каким будет оптимальное значение веса фонда в портфеле. И выбор веса фонда в соответствии с профилем риска инвестора на практике дает пусть и не оптимальное, но приемлемое значение доходности стратегии в целом.

Литература

1. Исаакман М. Как инвестировать в индексы. — М: Альпина, 2003.
2. Wild R. *Index Investing for Dummies*, Wiley, 2009.
3. Буренин А.Н. Управление портфелем ценных бумаг. — М: ВТО им. С. И. Вавилова, 2015.
4. Касимов Ю. Ф., Аль-Натор М. С., Колесников А. Н. Основы финансовых вычислений. Основные схемы расчета финансовых сделок. Учебник для вузов. — М.: Кнорус, 2017.
5. Касимов Ю. Ф., Аль-Натор М. С., Колесников А.Н. Основы финансовых вычислений. Оптимизация портфелей, облигации и хеджирование. Учебник для вузов. — М.: Кнорус, 2017.
6. Rasmussen M. *Quantitative Portfolio Optimization*, Palgrave, 2003.

UDC 519.864.3

Optimal Weights of Risk Asset in the Index Portfolio of Constant-Proportion Strategies

Y. F. Kasimov, M. I. Timerbaev

*Department of Data Analysis, Decision Making, and Financial Technology
Financial University under the Government of Russian Federation
49 Leningradsky Prospekt, Moscow, 125993, GSP-3, Russian Federation*

Email: y.f.kasimov@mail.ru, timerbayevgroup@yandex.ru

This paper analyzes the optimal of constant proportion index portfolio strategies. They are passive strategies and are becoming more common in Russia and abroad. The main problem in these strategies is the choice of the proportions in which the investor allocates his capital between a risky and risk-free assets. In constant proportion index portfolio the weight of risk asset remains constant throughout investment period. According on the daily value of units of these funds and the annual interest rate for the 11-year period, using a specially developed program optimal weight index funds in the portfolios has been found. For each Fund was determined the sequence of optimal weights and the corresponding optimum yield for consecutive investment periods with a specified frequency of re-balancing. It was found that in almost all cases, the optimal weights of fund equals the extreme values 0 or 1. Also, the frequencies of these values in the selected sequence is about the same for all funds.

Key words and phrases: passive investment strategy; constant proportion strategy, indices, index portfolios, the optimal index weights.

УДК 519.217.8

Управление марковской цепью с непрерывным ограниченным множеством состояний, оптимизирующее стационарное распределение цепи

М. Г. Коновалов*, Р. В. Разумчик*†

* *Институт проблем информатики
Федеральный исследовательский центр «Информатика и управление» РАН
ул. Вавилова, д. 44, к. 2, Москва, Россия, 119333*† *Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклухо-Маклая, д. 6, Москва, Россия, 117198*

Email: mkononov@ipiran.ru, rrazumchik@ipiran.ru, razumchik_rv@rudn.university

Рассматривается управляемая марковская цепь с непрерывным ограниченным множеством состояний. Ставится задача нахождения такой стратегии управления, при которой стационарное распределение цепи максимально приближено к некоторому заданному распределению на множестве состояний. Приводится мотивация постановки задачи и обсуждается ее решение.

Ключевые слова: цепь Маркова, адаптивное управление, непрерывное множество состояний, робототехника.

Рассмотрим цепь Маркова $\{x_n, n \geq 0\}$, принимающую значения из отрезка $[0, 1]$. Начальное распределение цепи несущественно, а переходная вероятность определяется следующим образом. Пусть задано число $\theta \in (0, 1]$. Выберем функцию $s(x)$, такую, что $0 \leq s(x) \leq 1$ при $x \in [0, 1]$ и условимся, что эволюция цепи определяется формулой

$$x_{n+1} = \begin{cases} x_n - \frac{\xi_n x_n}{\theta}, & \text{с вероятностью } s(x_n), \\ x_n + \frac{\xi_n (1-x_n)}{\theta}, & \text{с вероятностью } 1 - s(x_n), \end{cases} \quad n = 0, 1, \dots, \quad (1)$$

где ξ_n — взаимно независимые случайные величины, распределенные равномерно на $[0, 1]$. При определенных ограничениях на функцию $s(x)$ существует единственное стационарное, абсолютно непрерывное распределение цепи $\{x_n, n \geq 0\}$, не зависящее от начального состояния. Обозначая через $\pi(x)$ плотность этого распределения, можно убедиться, что она удовлетворяет следующему функциональному уравнению:

$$\pi(x) = \theta \int_{\max(0, \frac{\theta x - 1}{\theta - 1})}^x \pi(u) \frac{1 - s(u)}{1 - u} du + \theta \int_x^{\min(1, \frac{\theta x}{\theta - 1})} \pi(u) \frac{s(u)}{u} du.$$

Требуется найти функцию $s(x)$, которая минимизирует значение целевой функции

$$\int_0^1 (\pi(x) - \rho(x))^2 dx, \quad (2)$$

где $\rho(x)$ — некоторая заданная плотность распределения на отрезке $[0, 1]$. В предположении, что функция $s(x)$ принадлежит некоторому заданному множеству параметризованных функций, предлагается метод нахождения значений параметров, при которых целевая функция (2) принимает минимальное значение.

Побудительным мотивом к постановке задачи является следующий пример из [1, 2, 4–7]. Пусть по замкнутой и ограниченной области S , имеющей форму

прямоугольника размера $X \times Y$, передвигается робот. Двигается он дискретно-непрерывным образом: находясь в некоторой точке (x_n, y_n) , робот сначала выбирает следующую точку (x_{n+1}, y_{n+1}) для посещения, затем движется к ней по прямой линии¹, далее выбирает следующую точку (x_{n+2}, y_{n+2}) и т.д. Одной из задач робота является регулярное посещение каждого участка области S , причем некоторые заданные заранее части области S должны посещаться им чаще, чем другие. Для реализации этой цели будем предполагать, что робот имеет встроенные механизмы определения своего местоположения (например, датчики измерения расстояния). Для определенности примем (aX, bY) , $0 < a, b < 1$, за «точку притяжения». Окрустность этой точки робот может посещать чаще, чем другие подобласти S . Для выполнения этой задачи в робота необходимо «заложить» не зависящее от времени правило выбора очередной точки, в зависимости от его текущего местоположения (т.е. стационарную стратегию перемещения). Но не вполне ясно (а точнее, и вовсе не ясно), какая стратегия может обеспечить достижение поставленной цели: робот может «защиклиться» на какой-то области отличной от необходимой (т.е. от окрестности (aX, bY)), или вовсе не посещать некоторые точки области S .

К примеру, робот может двигаться по следующему (стационарному) правилу. Пусть (x_n, y_n) местоположение робота в после n -го по счету перемещения. Сначала робот выбирает направления движения: если $0 < x_n < aX$, то робот выбирает движение направо (\rightarrow), иначе — налево (\leftarrow); если $0 < y_n < bY$, то робот выбирает движение вперед (\uparrow), иначе — назад (\downarrow). Затем он рассчитывает координаты следующей точки по правилу

$$(x_{n+1}, y_{n+1}) = \begin{cases} (\xi_{1n} \times (1 - x_n), \xi_{2n} \times (1 - y_n)), & \text{если } (\rightarrow, \uparrow), \\ (\xi_{1n} \times x_n, \xi_{2n} \times (1 - y_n)), & \text{если } (\leftarrow, \uparrow), \\ (\xi_{1n} \times (1 - x_n), \xi_{2n} \times y_n), & \text{если } (\rightarrow, \downarrow), \\ (\xi_{1n} \times x_n, \xi_{2n} \times y_n), & \text{если } (\leftarrow, \downarrow), \end{cases} \quad (3)$$

где ξ_{1n}, ξ_{2n} — взаимно независимые равномерно распределенные на $[0, 1]$ случайные величины, и передвигается в нее. Далее он опять выбирает направления движения и переходит из точки (x_{n+1}, y_{n+1}) в точку (x_{n+2}, y_{n+2}) по правилу (3) и т.д. Из данного описания следует, что положение робота в последовательные моменты времени (сразу после очередного перемещения) есть цепь Маркова $\{z_n = (x_n, y_n), n \geq 0\}$ с непрерывным множеством состояний $[0, X] \times [0, Y]$, причем каждая из компонент есть цепь Маркова типа (1) с $\theta = 1$. Ввиду того, что направления передвижения каждый раз выбираются независимо, из (3) нетрудно выписать выражение для плотности вероятности перехода $p(x_{n+1}, y_{n+1} | x_n, y_n)$:

$$p(x_{n+1}, y_{n+1} | x_n, y_n) = \left(\mathbf{1}_{\{x_n < x_{n+1}\}} \frac{1 - s_a(x_n)}{1 - x_n} + \mathbf{1}_{\{x_n > x_{n+1}\}} \frac{s_a(x_n)}{x_n} \right) \times \\ \times \left(\mathbf{1}_{\{y_n < y_{n+1}\}} \frac{1 - s_b(y_n)}{1 - y_n} + \mathbf{1}_{\{y_n > y_{n+1}\}} \frac{s_b(y_n)}{y_n} \right),$$

где $\mathbf{1}_{\{A\}}$ — индикаторная функция множества A и

$$s_a(x) = \begin{cases} 0, & 0 \leq x \leq aX, \\ 1, & aX < x \leq 1, \end{cases}$$

¹ Другие характеристики движения (скорость, ускорение и пр.) значения для данного примера не имеют.

$$s_b(x) = \begin{cases} 0, & 0 \leq x \leq bY, \\ 1, & bY < x \leq 1. \end{cases}$$

При определенных ограничениях существует единственное стационарное распределение цепи $\{z_n, n \geq 0\}$, не зависящее от начального состояния. Обозначим плотность этого распределения через $p(x, y)$. В [1] показано, что цепи $\{x_n, n \geq 0\}$ и $\{y_n, n \geq 0\}$ являются эргодическими и, значит, цепь $\{z_n, n \geq 0\}$ является эргодической. Кроме того в [1] показано, что стационарная плотность $p(x, y)$ цепи $\{z_n, n \geq 0\}$ имеет глобальный максимум в точке (aX, bY) . Если же для робота заданы несколько «точек притяжения» или ограничения на доли времени, прерываемые в каждой подобласти S , то не понятно как можно модифицировать правило (3). Это связано с тем, что для осуществления модификации необходимо знать вид стационарной плотности $p(x, y)$ и, как можно понять из (1), за исключением простейших случаев стратегий, даже в одномерном случае получить для нее явный вид не представляется возможным. Таким образом, при передвижении робота в произвольной ограниченной области в соответствии с заранее заданными ограничениями на доли времени, проводимые роботом в каждой подобласти, возникает задача нахождения соответствующей стратегии управления.

В докладе предлагается способ нахождения стратегии, минимизирующей отклонение фактической стационарной плотности от заданного «эталона» для одномерного случайного блуждания. Модель объекта расширена по сравнению с [1, 2] привнесением параметра (θ) , ограничивающего амплитуду одного шага.

Предложенный для решения задачи алгоритм градиентного типа основан на выводе формулы для градиента целевой функции по параметрам стратегии и построении статистических оценок частных производных фактической стационарной плотности, основанных на наблюдениях состояний цепи. Алгоритм предназначен для работы с реальной (или имитируемой на компьютере) траекторией системы и относится к семейству адаптивных алгоритмов управления марковскими цепями [3]. Рассмотренная оптимизационная задача имеет существенные отличия от традиционного марковского процесса принятия решений (м.п.п.р.), и ее решение потребовало развития теории. Наиболее важный момент заключается в отсутствии в модели одношагового дохода. Дополнительные трудности связаны с выбором нестандартной для м.п.п.р. целевой функции, а также с наличием несчетного множества состояний цепи.

Литература

1. Ramli M. A., Leng G. The stationary probability density of a class of bounded Markov processes // *Adv. Appl. Probab.* — 2010. — Vol. 42. — Pp. 986–993.
2. McKinlay S., Borovkov K. On explicit form of the stationary distributions for a class of bounded Markov chains // *J. Appl. Probab.* — 2016. — Vol. 53 (1). — Pp. 231–243.
3. Коновалов М. Г. Методы адаптивной обработки информации и их приложения. — М.: ИПИ РАН, 2007.
4. Diaconis P., Freedman D. Iterated Random Functions *SIAM review.* — 1999. — Vol. 41, no. 1. — Pp. 45–76.
5. Forte B., Vrscay E. R. Solving the inverse problem for measures using iterated function systems: A new approach // *Journal of Applied Probability and Advances in Applied Probability.* — 1995. — Vol. 27. — Pp. 800–820.
6. Ladjimi F., Peigné M. Iterated function systems with place dependent probabilities and application to the Diaconis-Friedman's chain on $[0, 1]$. 2017. <hal-01567392>
7. Peigné M. Iterated Function Systems and Spectral Decomposition of the Associated Markov Operator // *Publications mathématiques et informatique de Rennes*, 1993. — Pp. 1–28.

UDC 519.217.8

Control of a Discrete-Time Markov Chain with Continuous Bounded State Space which Leads to the Given Stationary Distribution

M. G. Konovalov*, R. V. Razumchik*†

** Institute of Informatics Problems
Institute of Informatics Problems, FRC CSC RAS
44-2 Vavilova str., Moscow, 119333, Russian Federation*

*† Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: mkonvalov@ipiran.ru, rrazumchik@ipiran.ru, razumchik_rv@rudn.university

We consider a controlled Markov chain with discrete time and continuous bounded state space. The problem under consideration is to find such a control policy for the chain under which the stationary distribution of the chain is as close as possible to the given stationary distribution. Motivation for this problem is illustrated through the example.

Key words and phrases: Markov chain, adaptive control, continuous state space, robotics.

УДК 519.248.6

Характеристики неравномерности распределения ресурсов в экономических системах

А. В. Крянев*, Д. Е. Слива*, В. В. Матохин†

* *Национальный исследовательский ядерный университет «МИФИ»*† *ЗАО «ТЕКОРА»*

Email: avkryanev@mephi.ru, desliva@mephi.ru, vmatokhin@tekora.ru

В докладе исследуются свойства ранее представленного двухпараметрического семейства функций, аппроксимирующих кривые Лоренца. Проведен анализ различных функций плотности распределения, соответствующих различным значениям параметров семейства функций, аппроксимирующих кривую Лоренца. Исследована зависимость между такими статистическими характеристиками распределений, как величина межквартильного диапазона и коэффициент эксцесса, и параметрами рассматриваемого семейства функций. Предложено толкование экономического содержания значений параметров семейства функций, аппроксимирующей рассматриваемые кривые Лоренца.

Ключевые слова: распределение ресурсов, кривая Лоренца, дифференциация доходов.

1. Введение

Одним из возможных представлений распределения ресурсов внутри экономической системы является кривая Лоренца. В литературе предложено множество аппроксимаций кривой Лоренца, отчасти исследованы свойства аппроксимирующих ее функций и характеристик соответствующих функций распределения [1–6]. В данном докладе исследуются свойства одной из таких аппроксимаций с целью выделения особенных значений параметров, при которых соответствующее распределение ресурсов обладает полезными свойствами. Также предлагается рассматривать параметры функции, аппроксимирующей кривую Лоренца в качестве различных по содержанию характеристик неравномерности распределения ресурсов внутри системы.

2. Анализ характеристик распределения

Рассматривается распределение некоторого ресурса G между N экономическими агентами. Пусть $i = 1, \dots, n, \dots, N$ — индексы агентов, обладающих объемами ресурса G_i :

$$\sum_{i=1}^N G_i = G.$$

Распределение ресурсов G_i , $i = 1, \dots, N$ может быть представлено в виде диаграммы Лоренца [5, 6]. Для этого предварительно проводится ранжирование исходного исследуемого ряда:

$$G_1 \leq G_2 \leq \dots \leq G_N.$$

Затем рассчитываются накопленные суммы ресурсов:

$$S_n = \sum_{i=1}^n G_i$$

и полученные результаты представляются в виде графика на плоскости с осями $x = n/N$ и $y = S_n/S_N$.

В работах [1, 2] показано, что кривые Лоренца могут быть аппроксимированы двухпараметрической кривой из семейства функций

$$y(x, \alpha, \beta) = 1 - (1 - x^\alpha)^{1/\beta},$$

где $1 \leq \alpha < \infty$, $1 \leq \beta < \infty$.

Ранее было показано [3, 4], что параметры α и β выступают в качестве показателей степени неравномерности распределения ресурсов. При $\alpha = \beta = 1$ $y(x, \alpha, \beta) = x$, что соответствует совершенно равномерному распределению богатства, а при $\alpha \rightarrow \infty$, $\beta \rightarrow \infty$ распределение стремится к абсолютно неравномерному распределению.

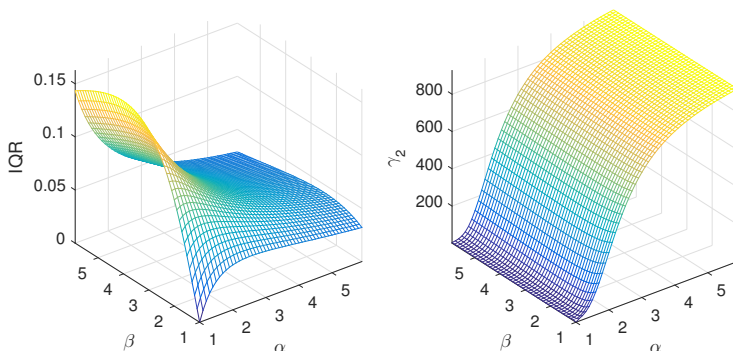


Рис. 1. Зависимость величины интерквартильного размаха и коэффициента эксцесса от значений показателей α и β

Утверждается, что параметры α и β характеризуют различные свойства эмпирических распределений доходов населения.

Параметр α характеризует острровершинность распределения, менее пологого при больших значениях параметра. При малых значениях α значения G_i сосредоточены вблизи среднего, а с возрастанием параметра значения G_i распределены далеко от него. Математически увеличение параметра α проявляется в возрастании коэффициента эксцесса распределения, рассчитываемого, как

$$\gamma_2 = \frac{\mu_4}{\sigma^4} - 3,$$

где $\mu_4 = \mathbf{E}[(G - \mathbf{E}G)^4]$ — четвертый центральный момент случайной величины G , а σ — её среднеквадратичное отклонение.

В свою очередь параметр β характеризует полярность общества, проявляющуюся в тенденции к выделению класса сверхбедных и сверхбогатых. Значение параметра определяет величину межквартильного размаха IQR , рассчитываемую как разность между медианой 50% самых богатых и медианой 50% самых бедных групп населения. Величина межквартильного диапазона является устойчивой мерой вариации, характеризующей разброс значений случайной величины. С экономической точки зрения величина межквартильного диапазона показывает разрыв в доходах между 25% самых богатых и 25% самых бедных групп населения.

На рис. 1 проиллюстрированы зависимости величины межквартильного размаха и коэффициента эксцесса от значений параметров α и β . По осям абсцисс и ординат

отложены значения показателей α и β , а по оси аппликат — значения исследуемых характеристик распределения.

Показано, что значение коэффициента эксцесса рассматриваемого семейства распределений определяется параметром α . При этом показатель β незначительно влияет на величину коэффициента эксцесса по сравнению с α . Можно видеть, что максимальное значение величины межквартильного диапазона достигается при больших значениях показателя β и $\alpha = 1$. Таким образом, можно заключить, что показатель α характеризует распределение богатства с точки зрения разрыва между максимальным и минимальным уровнем, а показатель β является мерой вариации.

3. Заключение

Полученные результаты могут быть применены в рамках решения задачи об оптимальном распределении ресурсов между несколькими объектами. Предложенная трактовка экономического содержания параметров α и β функции, аппроксимирующей кривую Лоренца, позволяет использовать их для выделения и анализа классов и кластеров экономических систем.

Литература

1. Antoniou I., Ivanov V.V., Korolev Yu.L. Kryanev A.V., Matokhin V.V., Suchanecki Z. Analysis of resources distribution in economics based on entropy // *Physica A*. — 2002. — No. 304. — Pp. 525–534.
2. Antoniou I., Ivanov V.V., Kryanev A.V., Matokhin V.V., Shapovalov M.V. On the efficient resources distribution in economics based on entropy // *Physica A*. — 2004. — No. 336. — Pp. 549–562.
3. Крянев А. В., Матохин В. В., Харитонов В. В. Энтропийный метод мониторинга реализации экономических систем // *Экономические стратегии*. — 2010. — No. 5.
4. Крянев А. В., Матохин В. В., Климанов С. Г. Статистические функции распределения ресурсов в экономике. — М.: Препринт МИФИ, 010-98, 1998.
5. Kleiber C., Kotz S. Statistical size distributions in economics and actuarial sciences. — John Wiley & Sons, 2003.
6. Modeling income distributions and Lorenz curves / Ed. by D. Chotikapanich. — Springer Science & Business Media, 2008.

UDC 519.248.6

Irregularity Measures of Resource Distributions within Economic Systems

A. V. Kryanev*, D. E. Sliva*, V. V. Matokhin†

* *National Research Nuclear University "MEPhI"*

† *ZAO "TEKORA"*

Email: avkryanev@mephi.ru, desliva@mephi.ru, vmatokhin@tekora.ru

In this report the research on two-parameter Lorenz curve approximants is continued. Distribution density functions that correspond to various values of parameters of the class of functions used to approximate the Lorenz curve are analysed. The ties between the values of the parameters and statistic estimators are tested. It is concluded that the values of the approximant parameters are connected to the coefficient of kurtosis and the value of the interquartile range. An interpretation of the approximant parameters with regards to economic inequality is suggested.

Key words and phrases: resource distribution, Lorenz curve, economic inequality.

УДК 519.622.1

Комплекс программ для решения нелинейных дифференциальных уравнений в частных производных первого порядка

Я. Ю. Кузив

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: yaroslav.kuziv@yandex.ru

Задача этой работы описать методы для решения дифференциальных уравнений в частных производных первого порядка. Эти методы будут использоваться для численного решения распространения высокочастотных акустических, упругих и электромагнитных волн, которые можно использовать во многих прикладных задачах.

Наша основная задача — написать комплекс программ для решения уравнения в частных производных в кусочно определенных областях. Для этого был разработан метод для решения уравнений, используемых в том числе и в геометрической оптике. Уравнения будем решать с помощью метода характеристик и метода Рунге–Кутты.

Помимо этого были написаны функции для нахождения конусов Коши и линии уровня, что дает дополнительную информацию для анализа полученных результатов и проверка корректности решения. Данные результаты помогут в дальнейших исследованиях задач геометрической физике.

Ключевые слова: характеристика, метод Рунге–Кутты, волновой фронт.

1. Введение

Общее нелинейное уравнение в частных производных первого порядка с двумя независимыми переменными имеет вид

$$F(x, y, u, p, q) = 0, \quad (1)$$

где $p = \frac{\partial u}{\partial x}$, $q = \frac{\partial u}{\partial y}$. Такие уравнения часто встречаются в аналитической механике, теории оптимального управления, динамическом программировании, геометрической оптике и во многих других областях. Например, мы можем решать уравнения геометрической оптике, такие как уравнение эйконала [1].

Уравнения будут решаться с помощью метода характеристик. Для этого надо свести уравнение к характеристической системе обыкновенных дифференциальных уравнений [2]

$$\frac{dx}{F_p} = \frac{dy}{F_q} = \frac{du}{p \cdot F_p + q \cdot F_q} = \frac{-dp}{F_x + p \cdot F_u} = \frac{-dq}{F_y + q \cdot F_u}, \quad (2)$$

где $F_x = \frac{\partial F}{\partial x}$, $F_y = \frac{\partial F}{\partial y}$, $F_u = \frac{\partial F}{\partial u}$, $F_p = \frac{\partial F}{\partial p}$, $F_q = \frac{\partial F}{\partial q}$.

Из (2) мы можем найти кривые, именуемые характеристиками. В начале, мы задаем начальные значения для каждой характеристики, определяющие его поведение в начальный момент времени, так чтобы (1) являлось тождеством. Можно задавать условия двумя способами: в виде “светящейся точки” и в виде кривой. В случае “светящейся точки” мы задаем переменные x, y, u , которые задают значение функции в начальный момент, а переменные p и q задают начальные скорости изменения u (углы под которыми выходят характеристики). Помимо этого можно задать условия в виде кривой. В данном случае каждая точка кривой задается с помощью переменных x, y, u и скоростями p и q , задающими угол, под которыми

характеристики будут двигаться в начальный момент. Так же надо следить, чтобы эти условия удовлетворяли (1).

Численное решение дифференциальных уравнении в частных производных первого порядка используется во многих областях (физика, химия, математическое моделирование и многие другие) и имеет большое количество применений. Поэтому написание комплекса программ для решения подобных уравнений является очень актуальной задачей. И существует несколько комплексов программ, которые могут решать подобные задачи. Например, метод характеристик реализован у Maple в пакете для решения дифференциальных уравнений. Но он позволяет решать уравнение, заданное единым символьным выражением во всем рассматриваемой области, т.е. мы не можем решать задачу, которая состоит из нескольких под областей, что встречается очень часто в геометрической оптике. Поэтому была выбрана эта тема для исследования и написания программы.

2. Пакет PDE для Python

Этот метод был реализован на объектно-ориентированном языке программирования Python в среде программирования Jupyter Notebook. Задача решается в пространстве R^5 , точками которого служат всевозможные наборы (x, y, z, p, q) , которое поделено на несколько областей G_1, \dots, G_k . Заданы сами эти области в виде неравенств G_1, \dots, G_k и символьные выражения для F_1, \dots, F_k , такие что

$$F|_{G_i} = F_i \quad \forall i = 1, \dots, k.$$

Для решения системы дифференциальных уравнений использовался метод Рунге–Кутты четвертого порядка. Используя вышеприведенные формулы и методы из уравнения $F(x, y, u, p, q) = 0$, мы можем получить набор характеристик, с помощью которых можно изобразить поверхность, описывающую частное решение.

3. Примеры

Рассмотрим в качестве примера уравнение эйконала $p^2 + q^2 = n^2(x, y)$, где n^2 имеет следующий вид:

$$n^2 = \begin{cases} 2 - x^2 - y^2, & x^2 + y^2 < 1 \\ 1, & x^2 + y^2 > 1 \end{cases}$$

В данном примере подразумевается, что мы имеем две области: круглую линзу Лüneбурга и среду в которой она находится (например воздух) [3]. Решение, полученное из набора характеристик, продемонстрировано на рис. 1.

Из полученных результатов мы можем найти линии уровня. Под линиями уровня мы будем понимать кривые в которых функция $u = C$, где $C = const$. Так же можно получить конусы Коши, которые позволяют нам судить о правильности решения. Т.е. если конус Коши будет касаться нашего решения, следовательно наше решение посчитано с достаточной точностью и его можно считать правильным.

4. Заключение

Представлено приложение для численного решения нелинейного дифференциальных уравнений в частных производных первого порядка, которую можно задать в пространстве, состоящем из нескольких под областей. Вычисление функции на нескольких под областях является новизной моей работы и будет использоваться в дальнейших исследованиях.

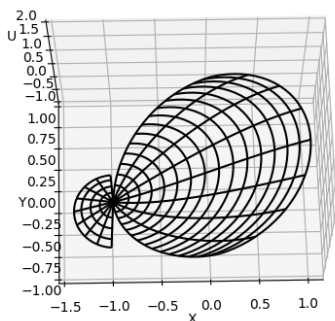


Рис. 1. Поверхность, описывающая частное решение уравнения эйконала

Литература

1. Enquist B., Runborg O. Computational high frequency wave propagation. — Cambridge University Press, 2003.
2. Зайцев В. Ф., Полянин А. Д., Справочник по дифференциальным уравнениям в частных производных первого порядка, Физико-математическая литература, 2003.
3. Lock J. Scattering of an Electromagnetic Plane Wave by a Luneburg Lens // Journal of the Optical Society of America A: Optics Image Science and Vision. — 2008. — Vol. 25. — Pp. 2980-2990. OCIS codes:080.2710, 080.5692.

UDC 519.622.1

The Complex of Programs for Solving Nonlinear Partial Differential Equations of the First Order

Ia. Iu. Kuziv

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: yaroslav.kuziv@yandex.ru

The task of this article is to describe methods for solving first-order partial differential equations. These methods will be used to numerically solve the propagation of high-frequency acoustic, elastic and electromagnetic waves, which can be used in many applied problems.

Our main task is to write a complex of programs for solving the partial differential equations in piecewise defined domains. For this purpose, a method for solving equations was used, including in geometric optics. The equations will be solved using the method of characteristics and the Runge-Kutta method.

In addition, functions for finding Cauchy cones and a level line were written, which provides additional information for analyzing the results obtained and verifying the correctness of the solution. These results will help in further investigations of problems in geometric physics.

Key words and phrases: characteristic, the Runge–Kutta method, wave front.

УДК 004.4

Численный анализ геометризованных уравнений Максвелла**Д. С. Кулябов*[†], М. Н. Геворкян*, А. В. Королькова**** *Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*[†] *Лаборатория информационных технологий
Объединённый институт ядерных исследований
ул. Жолио-Кюри, д. 6, Дубна, Московская область, Россия, 141980*

Email: kulyabov_ds@rudn.university, gevorkyan_mn@rudn.university, korolkova_av@rudn.university

В работе показано применение метода характеристик к уравнению эйконала, записанного для линз Максвелла и Люнеберга. Уравнение эйконала преобразовано к системе из четырех ОДУ. Полученная система решена стандартными численными методами. Описаны детали подхода к численному моделированию и построены изображения траекторий лучей и фронтов волн.

Ключевые слова: уравнение эйконала; линза Люнеберга; линза Максвелла; метод характеристик; Julia.

1. Введение

Уравнение эйконала связывает волновой и геометрических подход к описанию оптических явлений. Оно выводится из уравнений Максвелла без токов и зарядов и при условии гармонических электромагнитных волн вне проводящей изотропной среде [1]. Будем рассматривать проекцию лучей на плоскость Oxy и в этом случае уравнение эйконала сводится к двумерному виду:

$$\begin{cases} \left(\frac{\partial u(x, y)}{\partial x}\right)^2 + \left(\frac{\partial u(x, y)}{\partial y}\right)^2 = n^2(x, y), & (x, y) \in \mathbb{R}^2, \\ u(x, y) = \varphi(x, y), & (x, y) \in \mathbb{R}^2, \end{cases}$$

где $\mathbf{r} = (x, y, z)^T$ — радиус-вектор, $\varphi(\mathbf{r})$ — граничное условие, $n(\mathbf{r})$ — показатель преломления среды. Функция $u(\mathbf{r})$ является вещественной скалярной функцией и имеет физический смысл времени, ее часто называют *эйконалом*.

Для решения данного уравнения применим метод характеристик [2–4]. В результате уравнение в частных производных первого порядка сведется к системе из четырех обыкновенных дифференциальных уравнений первого порядка путем замены $p_1 = \frac{\partial u}{\partial x}$, $p_2 = \frac{\partial u}{\partial y}$:

$$\begin{cases} \frac{dx}{dt} = \frac{p_1}{n^2}, & \text{начальные условия} \\ \frac{dy}{dt} = \frac{p_2}{n^2}, & x(t)|_{t=0} = x_0, \\ \frac{dp_1}{dt} = \frac{1}{n} \frac{\partial n}{\partial x}, & y(t)|_{t=0} = y_0, \\ \frac{dp_2}{dt} = \frac{1}{n} \frac{\partial n}{\partial y}. & p_1(t)|_{t=0} = c_1 n(x_0, y_0), \\ & p_2(t)|_{t=0} = c_2 n(x_0, y_0). \end{cases}$$

Рассмотрим теперь показатели преломления среды для линз Люнеберга и Максвелла и опишем особенности численного моделирования этих линз.

2. Линза Люниберга

Линза Люниберга представляет собой сферическую линзу радиуса R с центром в точке (X_0, Y_0) (рассматриваем проекцию на плоскость Oxy) с коэффициентом преломления следующего вида:

$$n(x, y) = \begin{cases} n_0 \sqrt{2 - \left(\frac{r}{R}\right)^2}, & r \leq R, \\ n_0, & r > R, \end{cases}$$

где $r(x, y) = \sqrt{(x - X_0)^2 + (y - Y_0)^2}$ — расстояние от центра линзы до произвольной точки (x, y) плоскости. Из формулы следует, что коэффициент n непрерывно меняется от $n_0\sqrt{2}$ до n_0 , начиная от центра линзы и заканчивая ее границей. Коэффициент преломления среды вне линзы постояен и равен n_0 . Обычно полагают $n_0 = 1$.

3. Линза Максвелла

Линза Максвелла также представляет собой сферическую линзу радиуса R с центром в точке (X_0, Y_0) (рассматриваем проекцию на плоскость Oxy) с коэффициентом преломления следующего вида:

$$n(x, y) = \begin{cases} \frac{n_0}{1 + \left(\frac{r}{R}\right)^2}, & r \leq R, \\ n_0, & r > R. \end{cases}$$

4. Численное моделирование

Численное моделирование уравнений для линз Максвелла (рис. 1) и Люнеберга (рис. 2) было осуществлено с помощью классических методов Рунге–Кутты с постоянным шагом сетки. Выбор методов именно с постоянным шагом обоснован необходимостью построения фронтов волн. Для этого необходимо иметь значения точек траекторий разных лучей для одних и тех же моментов времени, что невозможно в случае применения методов с коррекцией шага (даже в случае полной выдачи).

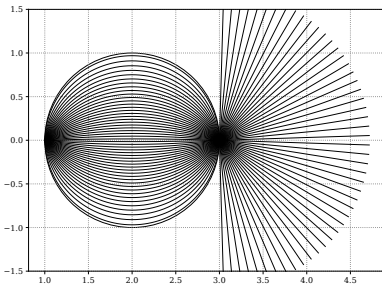


Рис. 1. Ход лучей внутри линзы Максвелла при точечном источнике и $n_0 = 1$

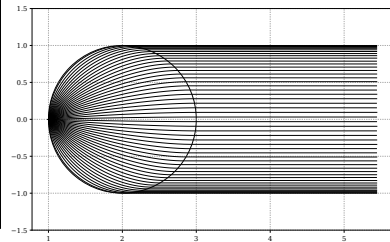


Рис. 2. Ход лучей внутри линзы Люнеберга при точечном источнике и $n_0 = 1$

5. Заключение

В работе представлена программная реализация решения уравнения эйконала для случая линз Лунеберга и Максвелла. Результаты визуализированны в виде траекторий лучей, проходящих через линзы и в виде фронтов электромагнитных волн.

Благодарности

Публикация подготовлена при поддержке программы РУДН «5-100» и при финансовой поддержке РФФИ в рамках научного проекта РФФИ 16-07-00556.

Литература

1. Борн М., Вольф Э. Основы оптики. — Москва : Наука, 1973.
2. Иванов Д. И., Иванов И. Э., Крюков И. А. Алгоритмы приближенного решения некоторых задач прикладной геометрии основанные на уравнении типа Гамильтона-Якоби // Журнал вычислительной математики и математической физики. — 2005. — Т. 45, № 8. — С. 1345–1358.
3. Кабанихин С. И., Криворотько О. И. Численное решение уравнения эйконала // Сибирские электронные математические известия. — 2013. — С. 28–34.
4. Москаленский Е. Д. О нахождении точных решений двумерного уравнения эйконала // Сибирский журнал вычислительной физики. — 2009. — Т. 12, № 2. — С. 201–209.

UDC 004.4

Numerical Analysis of Geometrized Maxwell Equations

D. S. Kulyabov*[†], M. N. Gevorkyan*, A. V. Korolkova*

* *Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

[†] *Laboratory of Information Technologies
Joint Institute for Nuclear Research
6 Joliot-Curie, Dubna, Moscow region, 141980, Russian Federation*

Email: kulyabov_ds@rudn.university, gevorkyan_mn@rudn.university, korolkova_av@rudn.university

This paper shows the application of geometric methods of electrodynamics to the calculation of optical devices, such as lenses Maxwell and Luneberg. The eikonal equation, which was transformed to the ODE system by the method of characteristics, is considered. The resulting system is written for the case of Maxwell and Luneberg lenses and solved by standard numerical methods. Describes the implementation details and images of the trajectories of rays and fronts of the waves.

Key words and phrases: eikonal equation; Luneberg lens; Maxwell lens; characteristics method; Julia.

УДК 519.6

Об одной обратной задаче для метагармонического уравнения

П. А. Лесик, А. А. Романов, Д. А. Чистотина

*Российский университет дружбы народов
ул. Миклухо-Маклая, д. 6, Москва, Россия, 117198*

Email: polinalesik@yandex.ru, anlromanov@gmail.ru, daria.chistotina97@mail.ru

Получено устойчивое приближенное решение обратной задачи к задаче Дирихле для метагармонического уравнения. Для его построения использован метод регуляризации Тихонова.

Ключевые слова: некорректно поставленная задача, обратная задача, метод регуляризации Тихонова.

1. Введение

Здесь мы рассматриваем задачу продолжения решения метагармонического уравнения с границы как задачу, обратную к задаче Дирихле для метагармонического уравнения аналогично задаче продолжения поля гармонического потенциала [1]. Задача некорректно поставлена, сведена к интегральному уравнению и для построения ее устойчивого приближенного решения используется метод регуляризации Тихонова [2].

2. Постановка задачи

Рассмотрим следующую краевую задачу — задачу Дирихле — для метагармонического уравнения:

$$D = \{(x, y, z) : -l_x < x < l_x, -l_y < y < l_y, -\infty < z < H\},$$

$$(\Delta - k^2) * \varphi(M) = 0, \quad M \in D,$$

$$\varphi|_{x=0, l_x} = 0; \quad \varphi|_{(y=0, l_y)} = 0,$$

$$\varphi|_{z=H} = \varphi_H,$$

$$\varphi \rightarrow 0 \text{ при } z \rightarrow -\infty.$$

Если функция φ_H известна, то решение задачи существует и единственно, а также устойчиво в соответствующих классах функций, то есть — корректно. В частности, можно получить нормальную производную этого решения на плоскости $z = 0$. Поставим обратную задачу. Пусть известна производная

$$\left. \frac{\partial \varphi}{\partial z} \right|_{z=0} = E_z^0,$$

$a \varphi_H$ — искомая функция. Поставим задачу восстановления функции φ_H в плоскости $z=H$ по функции E_z^0 , а, следовательно, и решения φ в области D . Эта задача сведена к интегральному уравнению Фредгольма первого рода:

$$\int_{\Pi(H)} K(M, P) \varphi_H(P) d\sigma_P = E_z^0(M), \quad M \in \Pi(0),$$

где $K(M, P)$ — ядро интегрального оператора — известная непрерывная функция.

Интегральное уравнение первого рода — некорректно поставленная задача. В качестве приближенного решения интегрального уравнения выбираем экстремаль функционала Тихонова:

$$M^\alpha[w] = \|Kwd\sigma - E_z^0\|_{L_2(\Pi(0))}^2 + \alpha\|w\|_{L_2(\Pi(H))}^2.$$

Экстремаль получена как решение уравнения Эйлера для функционала $M^\alpha[w]$ методом Фурье.

3. Выводы

Доказана сходимость приближенного решения к точному обратной задачи при стремлении к нулю погрешности во входных данных $E_z^0(M)$.

Литература

1. Лансеев Е. Б. Некорректные задачи продолжения гармонических функций и потенциальных полей и методы их решения. — М. Изд-во РУДН, 2006.
2. Тихонов А. Н., Арсенин В. Я. Методы решения некорректных задач. — М.: Наука, 1979.

UDC 519.6

Inverse Problem for the Equation Managementscope

P. A. Lesik, A. A. Romanov, D. A. Chistotina

*Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: polinalesik@yandex.ru, anromanov@gmail.ru, daria.chistotina97@mail.ru

The decision of one of the variants of the linear inverse problem of the potential for infinitely thin body, placed in a bilayer environment. For constructing stable approximate solution is used the regularization method of Tikhonov.

Key words and phrases: ill-posed problem, inverse problem, method of Tikhonov regularization.

УДК 519.6

Задача продолжения ньютоновского потенциала в цилиндре кругового сечения

Н. М. Лижненко, Д. А. Морозова

*Российский университет дружбы народов
ул. Миклухо-Маклая, д. 6, Москва, Россия, 117198*

Email: lizhnenko@live.ru, dasha_m96@mail.ru

Получено устойчивое приближенное решение некорректно поставленной краевой задачи для уравнения Лапласа. Для его построения использован метод регуляризации Тихонова.

Ключевые слова: некорректно поставленная задача, обратная задача, метод регуляризации Тихонова.

1. Введение

Рассматривается задача продолжения решения уравнения Лапласа в цилиндрическую область кругового сечения с границы, на которой это решение задано как след ньютоновского потенциала. Задача аналогична задаче продолжения поля гармонического потенциала в цилиндрическую область прямоугольного сечения [1]. Задача некорректно поставлена, сведена к интегральному уравнению Фредгольма первого рода и для построения ее устойчивого приближенного решения используется метод регуляризации Тихонова [2].

2. Постановка задачи

В области

$$D = \{(r, \varphi, z) : 0 < r < a, 0 < \varphi < 2\pi, -\infty < z < H\}$$

рассмотрим следующую краевую задачу — задачу Дирихле — для уравнения Лапласа:

$$\begin{aligned} \Delta u(M) &= 0, & M \in D, \\ u|_{r=a} &= 0, & u|_{\varphi=0} = u|_{\varphi=2\pi}, \\ u|_{z=H} &= u_H, \\ u &\rightarrow 0 & \text{при } z \rightarrow -\infty. \end{aligned}$$

Если функция u_H известна, то решение задачи существует и единственно, и может быть получено в явном виде — в виде разложения в ряд по функциям Бесселя. Это решение устойчиво в соответствующих классах функций.

Пусть теперь функция u_H не известна, но известна функция u на плоскости $z = 0$

$$u|_{z=0} = u_0,$$

представляющая собой, например, след ньютоновского потенциала. Поставим обратную задачу восстановления функции u_H в плоскости $z = H$ по функции u_0 , заданной на плоскости $z = 0$, и решения u в области D .

Эта задача сводится к интегральному уравнению Фредгольма первого рода:

$$\int_{\Pi(H)} K(M, P) u_H(P) d\sigma_P = u_0, \quad M \in \Pi(0).$$

где $K(M, P)$ — ядро интегрального оператора — известная непрерывная функция. Интегральное уравнение первого рода — некорректно поставленная задача. В качестве приближенного решения интегрального уравнения будем рассматривать экстремаль функционала Тихонова:

$$M^\alpha[w] = \left\| \int_{\Pi} K w d\sigma - u_0 \right\|_{L_2(\Pi(0))}^2 + \alpha \|w\|_{L_2(\Pi(H))}^2.$$

Экстремаль получена как решение уравнения Эйлера для функционала $M^\alpha[w]$ методом Фурье.

3. Выводы

Доказана сходимостъ приближенного решения к точному решению обратной задачи при стремлении к нулю погрешности в задании функции u_0 .

Литература

1. Ланев Е. Б., Некорректные задачи продолжения гармонических функций и потенциалных полей и методы их решения. — М. Изд-во РУДН, 2006.
2. Тихонов А. Н., Арсенин В. Я., Методы решения некорректных задач. — М.: Наука, 1979.

UDC 519.6

The Problem of Continuation of the Newtonian Potential in a Cylinder of Circular Cross Section

N. M. Lizhnenko, D. A. Morozova

*Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: lizhnenko@live.ru, dasha_m96@mail.ru

A stable approximate solution of an ill-posed boundary value problem for the Laplace equation is obtained. To construct it, Tikhonov's regularization method was used.

Key words and phrases: ill-posed problem, inverse problem, method of Tikhonov regularization.

УДК 519.862

Повышение пенсионного возраста и решение проблемы снижения численности занятого населения

Ю. А. Мысина

*Департамент анализа данных, принятия решений и финансовых технологий
Финансовый университет при Правительстве РФ
Ленинградский просп., д. 49, Москва, Россия, 125993, ГСП-3*

Email: yulya_mysina@mail.ru

Построены и проанализированы модели доли трудоспособного населения и доли занятых в экономике. Смоделированы доли трудоспособного населения и доли занятых в экономике в случае повышения пенсионного возраста мужчинам и женщинам на 5 лет.

Ключевые слова: трудоспособное население, занятость, демографические волны.

1. Введение

Прогнозирование уровня занятости является обязательной задачей современного рынка занятости. Зная долю занятого населения, можно принять меры по обеспечению граждан как можно большим количеством мест, достойными условиями труда и заработной платой, что в свою очередь, повысит уровень жизни населения и снизит безработицу. Цель работы — моделирование и прогнозирование занятости с учетом демографических волн.

Результаты работы могут служить ориентиром, как изменится доля занятых в экономике при повышении пенсионного возраста [1], будет ли оно эффективным с точки зрения улучшения ситуации на рынке труда.

2. Построение модели доли трудоспособного и занятого населения

В работе проведен анализ динамики доли трудоспособного населения за период 1959–2013 гг. и построена модель с использованием сплайн функций и лаговых переменных [2], вид модели представлен ниже:

$$y(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 d_1 + \alpha_4 t d_1 + \alpha_5 t^2 d_1 + \alpha_6 y_{t-1} + \alpha_7 y_{t-2},$$

где $t = 1, \dots, 55$, $d_1 = 1$ при $0 < t \leq 36$ (первая волна трудоспособности) и $d_1 = 0$ при $t > 36$ (вторая волна трудоспособности).

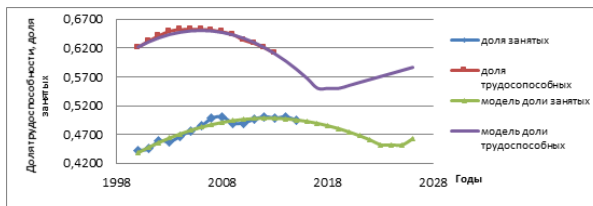


Рис. 1. Модель доли трудоспособного и занятого населения. Данные [4], расчеты автора

При этом минимум трудоспособности будет достигнут в 2017 году, и лишь в 2020 г. начнется постепенное увеличение за счет образования новой волны. Прогноз доли занятых в экономике был построен исходя из поведения доли трудоспособных: снижение занятости будет наблюдаться до 2023 года, и лишь в 2026 г. начнется постепенное увеличение (см. рис. 1).

Аналогичным образом были построены модели и спрогнозированы доли трудоспособных и занятых с учетом повышения пенсионного возраста мужчин — до 65 лет, женщин — 60 лет (рис. 2). В данном случае доля занятых в экономике увеличивается за счет добавление еще одной возрастной группы [3]

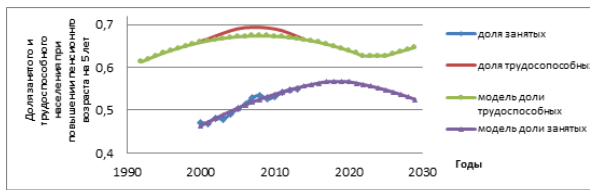


Рис. 2. Модель доли трудоспособного и занятого населения с учетом повышения пенсионного возраста мужчин до 65 лет, женщин — 60 лет

3. Заключение

В работе рассмотрены две ситуации развития занятости в России: первая — при общеустановленном пенсионном возрасте, вторая — при повышении пенсионного возраста на 5 лет. Отмечено, что повышение пенсионного возраста повысит долю занятых за счет добавления возрастной группы, которая будет постоянно увеличиваться за счет старения населения (см. табл. 1. В ней ТС — трудоспособный возраст). Однако стоит учитывать, что средняя продолжительность жизни мужчин в 2014 году составила лишь 65,3 года, а женщин — 76,5 лет. Это значит, что многие мужчины не доживут до пенсионного возраста [1]. В настоящее время в России имеется большая проблема с безработицей среди молодежи (см. табл. 2).

Таблица 1
Возрастная структура населения за период с 1989 по 2014 гг, тыс. человек

Год	Моложе ТС	В ТС	Старше ТС
1989	35995	83746	27196
2002	26327	88942	29778
2010	23126	89983	31714
2014	24717	85162	33788

При повышении пенсионного возраста данная проблема только усугубится: будет все меньше рабочих мест для молодых специалистов. Кроме этого, снизится производительность труда. Таким образом, на современном этапе развития России повышение пенсионного возраста не решит проблемы низкой занятости и увеличения доли трудоспособности. Огромное влияние на показатели рынка труда оказывают демографические волны, это необходимо учитывать при разработке мер по снижению безработицы и повышению занятости.

Таблица 2

Уровень безработицы по возрастным группам и виду населения в 2015 году,
%

Возраст	Городское население	Сельское население
15–19 лет	26	31,1
20–24 года	12,8	18
25–29 лет	5,7	10
30–39 лет	4,1	7,4
40–49 лет	3,1	6,6
50–59 лет	3,5	6,1
60–72 года	2,8	2,9

Благодарности

Автор признателен доценту Аль-Натору М. С., под научным руководством которого была выполнена данная работа.

Литература

1. Аль-Натор М. С., Донцова С. А., Соловьев А. К. Повышение пенсионного возраста в условиях социально-экономической неопределенности // Экономика и управление: проблемы, решения. — 2017. — Т. 3, № 6. — С. 69–73.
2. Тихомиров Н. П., Тихомирова Т. М., Ушмаев О. С. Методы эконометрики и многомерного статистического анализа. Учебник. — М.: Экономика, 2011.
3. Соловьев А. К. Исследование влияния неформальной занятости на сбалансированность бюджета ПФР и формирование пенсионных прав работника // Проблемы теории и практики управления. — 2017. — № 10. — С. 55–65.
4. Федеральная служба государственной статистики [Электронный ресурс]. — Ресурс доступа: <http://www.gks.ru>.

UDC 519.862

Rising of Retirement Age and Deciding the Problem of Decreasing the Amount of Employed Population

J. A. Mysina

*Department of Data Analysis, Decision Making, and Financial Technology
Financial University under the Government of Russian Federation
49 Leningradsky Prospekt, GSP-3, Moscow, 125993, Russian Federation*

Email: yulya_mysina@mail.ru

The share of able-bodied and employed population were modeled. On the assumption of an increase of the retirement age of men under 65 years, women- 60 years, the indexes were modeled again.

Key words and phrases: able-bodied population, employment, demographic waves.

УДК 519.864.3

Рыночные риски коммерческого банка: методы оценки и управления

Ю. С. Пермякова

*Департамент анализа данных, принятия решений и финансовых технологий
Финансовый университет при Правительстве РФ
Ленинградский просп., д. 49, Москва, Россия, 125993, ГСП-3*

Email: jusenka2013@gmail.com

Произведено тестирование методов оценки рыночных рисков: оценка стоимости под риском (VaR) и процедура бэкстестинга на примере облигационного портфеля АО АКБ «Международный финансовый клуб». Согласно полученным данным по портфелю ценных бумаг, в работе рассчитаны доходность облигаций портфеля, волатильность этих доходностей, позиции облигаций, текущий доход (расход) по портфелю. Установлено нарушение прогнозного значения модели, что в полной мере демонстрирует работу применяемых методов.

Ключевые слова: оценка стоимости под риском (VaR), бэкстестинг, портфель ценных бумаг, доходность, волатильность.

1. Введение

Данное исследование посвящено применению методов и инструментов управления рыночными рисками коммерческого банка, а также обоснованию направлений их практической реализации [1]. Предметом исследования являются теоретические и методические положения, определяющие рыночные риски коммерческого банка [2], а также механизм их оценки и управления: оценка стоимости под риском (VaR) и процедура бэкстестинга на примере конкретного коммерческого банка. Объектом исследования выступает измененный (в целях конфиденциальности) портфель ценных бумаг АО АКБ «Международный финансовый клуб».

2. Реализация методов оценки и управления

Для получения адекватной оценки VaR должны использоваться данные за последние 250 торговых дней (в случае наличия), но не менее 150 торговых дней. Текущая справедливая стоимость финансовых инструментов определяются по МСФО в соответствии с внутренними документами банка, регламентирующими, в том числе, оценку финансовых инструментов в отсутствие активного рынка [3, 4].

Расчет VaR портфеля финансовых инструментов производится по формуле:

$$\text{VaR}_{p,t+1} = a \sqrt{\sum_i \alpha_{i,t}^2 \sigma_{i,t}^2 + 2 \sum_j \sum_{i < j} \alpha_{i,t} \alpha_{j,t} \rho_{ij,t} \sigma_{i,t} \sigma_{j,t}},$$

где $\sigma_{i,t}$ — волатильность доходности финансового инструмента i на дату t , $\alpha_{i,t}$ — позиция по финансовому инструменту i на дату t , a — квантиль нормального распределения для вероятности 99%, $\rho_{ij,t}$ — коэффициент корреляции доходностей финансовых инструментов.

При проведении процедуры бэкстестинга используются прогнозные значения VaR и нереализованная прибыль или убыток (далее — P&L) по портфелю, рассчитываемый по формуле:

$$\text{P\&L}_t = \sum_i (P_{i,t} - P_{i,t-1}) \cdot N_{i,t-1} \cdot \text{Nom}_i \cdot \text{ExchRate}_{i,t-1},$$

где $P_{i,t}$ — текущая справедливая стоимость финансового инструмента i на дату t , Nom_i — номинал финансового инструмента i , $N_{i,t-1}$ — количество единиц финансового инструмента i в портфеле по состоянию на дату $t-1$, $ExchRate_{i,t-1}$ — курс Банка России, установленный на дату $t-1$. Нарушением прогноза модели считается случай, при котором нереализованная прибыль или убыток за один день превышает соответствующее прогнозное значение:

$$VaR_t < |P\&L_t|.$$

VaR, P&L, бэкестинг

Таблица 1

Дата	Var	P&L	Backtesting
31.12.2016	39 888 701	22 733 876	ok
09.01.2017	38 977 432	13 842 616	ok
...
02.03.2017	24 526 267	-4 236 987	ok
03.03.2017	24 485 002	2 142 763	ok
06.03.2017	24 637 806	-2 324 849	ok
07.03.2017	24 221 649	-27 963 556	breach
09.03.2017	24 213 554	-8 926 512	ok
10.03.2017	24 384 892	-7 517 338	ok
...

По произведенным расчетам можем заметить нарушение прогноза модели, произошедшее 7 марта 2017 года. P&L превысил по абсолютной величине прогнозное значение VaR (табл. 1). Это связано с тем, что цена на нефть марки Brent опустилась во время торгов 9 марта ниже 52 долларов за баррель. Такое сильное падение произошло впервые с 1 декабря 2017 года. Так, по состоянию на 14:00 по московскому времени стоимость майских фьючерсов на нефть марки Brent опускаясь до 51,62 доллара за баррель. В связи с тем, что произошло единственное нарушение, результаты бэкестинга следует отнести к «зеленой зоне», допускающей не более 4 превышений. Продемонстрируем это на графике (рис. 1).

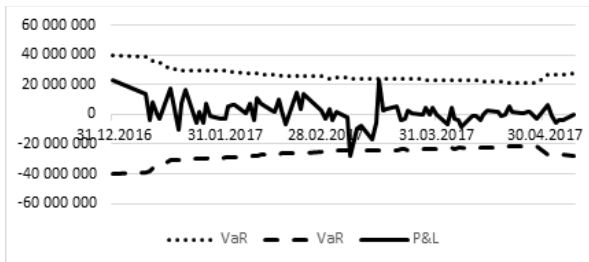


Рис. 1. Результаты бэкестинга

3. Заключение

В работе представлены теоретические основы анализа рыночных рисков коммерческих банков, их сущность и содержание. Протестированы методы оценки рыночных рисков: оценка стоимости под риском (VaR) и процедура бэкстестинга на примере облигационного портфеля АО АКБ «Международный финансовый клуб».

В ходе исследования методов оценки рыночных рисков было установлено нарушение прогнозного значения модели, что в полной мере продемонстрировало работу применяемых методов.

Благодарности

Автор признателен доценту Аль-Натор С.В., под научным руководством которой была выполнена данная работа.

Литература

1. Положение Банка России от 3 декабря 2015 г. № 511-П «О порядке расчета кредитными организациями величины рыночного риска».
2. Письмо Банка России от 23.06.2004 № 70-Т «О типичных банковских рисках».
3. Банковский менеджмент. Учебник / кол. авторов; под редакцией д-ра наук, проф. О.И. Лаврушина. — 2-е изд., перераб. и доп. — М.: КНОРУС, 2009.
4. Климова О. А. Организационные инструменты управления рыночными рисками // IV Международная научно-практическая конференция студентов и молодых ученых «Молодые экономисты – будущему России». Направление «Деньги, кредит и банковская деятельность». — Ставрополь: СКФУ, 2012.

UDC 519.864.3

Market Risks of the Commercial Bank: Methods of Estimation and Management

J. S. Permyakova

*Department of Data Analysis, Decision Making, and Financial Technology
Financial University under the Government of Russian Federation
49 Leningradsky Prospekt, GSP-3, Moscow, 125993, Russian Federation*

Email: jusenka2013@gmail.com

Methods for assessing market risks: valuation at risk (VaR) and backtesting procedure on the example of the bond portfolio of JSC JSCB International Financial Club (IFC Bank) have been tested. According to the received data on the securities portfolio, the profitability of the portfolio bonds, the volatility of these ones, the position of the bonds, the current income (expenditure) for the portfolio were calculated. A violation of the predicted value of the model is established, which fully demonstrates the work of the methods used.

Key words and phrases: valuation under risk (VaR), backtesting, securities portfolio, profitability, volatility.

УДК 004.9

Метод генерации случайных векторов в заданными вероятностными характеристиками

А. Ю. Попков*[†]** Институт системного анализа ФИЦ ИУ РАН
ул. Вавилова д. 44, к. 2, Москва, Россия, 119333**† Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: apopkov@isa.ru

Предлагается метод генерации случайных векторов с заданной функцией плотности распределения вероятностей, обладающий гарантированными оценками качества получаемых случайных векторов. Проведено экспериментальное исследование метода и его сравнение с алгоритмом Метрополиса–Гастингса.

Ключевые слова: генерация случайных векторов; метод Метрополиса–Гастингса.

1. Введение

Проблема генерации случайных величин с заданными вероятностными характеристиками является очень актуальной в различных прикладных областях. К таким областям, в частности, относится область байесовских методов, вычисление многомерных интегралов и др. Интерес к подобным методам резко возрос в последнее десятилетие в связи с качественным развитием вычислительной техники.

Методы генерации случайных векторов не единичной размерности, имеющих плотность распределения вероятностей (ПРВ) произвольной формы, насчитывают более чем 50-летнюю историю [1–3]. Широко используемые в настоящее время методы концептуально основаны на идее Acceptance-Rejection [1] и ее развитии Markov Chain Monte Carlo (MCMC) [2].

Метод Acceptance-Rejection (AR-метод) [1, 3] состоит в следующем. Для заданной, в общем случае, многомерной, ПРВ $w(\mathbf{x})$, определенной на компактном носителе \mathcal{G} , последовательно вычисляется случайная величина γ , равномерно распределенная на интервале $[0, \max w(\mathbf{x})]$, и случайный вектор η , равномерно распределенный на множестве \mathcal{G} . Если $\gamma \leq w(\eta)$, то η принимается в качестве реализации случайной величины с требуемым распределением. В противном случае процесс повторяется. Из этой простой конструкции видно, что метод обладает несколькими недостатками. Во-первых, время получения требуемого количества случайных чисел — случайно, и оно сильно зависит от конфигурации множества \mathcal{G} и функции w . Во-вторых, необходимо знать величину максимума плотности, что само по себе очень нетривиально в многомерном случае.

Наиболее распространенным сейчас является метод Метрополиса–Гастингса (МН-метод) [2]. В данном методе, каждый новый элемент случайной последовательности генерируется особым образом функционирующей марковской цепью. Эффективность МН-метода зависит от формы вспомогательного распределения, используемого для построения такой цепи. На практике в качестве такого распределения часто используется нормальное распределение, с предварительной настройкой его параметров так, чтобы лучше аппроксимировать заданную плотность [4–7]. Основной недостаток этого метода состоит в том, что неизвестны какие-либо оценки скорости сходимости распределения построенной марковской цепи к стационарному, то есть, к требуемому; более того, неизвестны даже условия существования стационарного распределения. Таким образом, на практике, приходится принимать полученные случайные числа в качестве удовлетворяющих требуемому распределению после применения некоторых эмпирических подходов, направленных на практическое решение указанной проблемы. Строгих теоретических обоснований такого подхода по-видимому не существует.

2. Постановка задачи и ее решение

Задача генерации случайных векторов с заданными вероятностными характеристиками состоит в следующем. Пусть задана ПРВ $w(\mathbf{x})$, где $\mathbf{x} \in \mathcal{G} \subset R^n$. Необходимо сгенерировать последовательность $\{\mathbf{x}_k\}$ случайных векторов \mathbf{x} , ПРВ которых были бы близки, в терминах принятого критерия близости, к заданной $w(\mathbf{x})$. В настоящей работе предполагается, что множество $\mathcal{G} = \Pi$, где Π — n -мерный параллелепипед.

Решение этой задачи состоит из двух основных этапов: *аппроксимации* заданной непрерывной плотности $w(\mathbf{x})$ кусочно-постоянной $\hat{w}(\mathbf{x})$, и последующей *генерации* из этой плотности с помощью генератора случайных чисел с равномерным распределением, доступного на любой вычислительной системе.

Для реализации аппроксимации в начале определяется оптимальный по каждой координате шаг h_k , $k = \overline{1, n}$ сетки, с помощью которой множество Π разбивается на N элементарных параллелепипедов Π_i с объемами v . Объемы всех параллелепипедов Π_i одинаковые и равны

$$v = \prod_{k=1}^n h_k, \quad N = V/v.$$

Качество аппроксимации предлагается измерять функционалом

$$J = \left| \int_{\mathcal{G}} (w(\mathbf{x}) - \hat{w}(\mathbf{x})) d\mathbf{x} \right|,$$

который характеризует расхождение между заданной и аппроксимированной вероятностями произвольного события.

Затем предлагается заменить функцию $w(\mathbf{x})$ на каждом элементарном параллелепипеде Π_i на оптимальную константами B_i^* .

Имея такую кусочно-линейную аппроксимацию, нетрудно сэмплировать из нее требуемые случайные векторы. Для этого, с помощью случайной величины β , принимающей целочисленные значения $1, \dots, N$ с вероятностями $p_i = vB_i^*$, производится выбор одного из элементарных параллелепипедов Π_i .

На каждом параллелепипеде Π_i генерируются случайные векторы ξ_i , не зависящие от β и равномерно распределенные на Π_i . ПРВ вектора ξ_i на Π_i равна $1/v$, а вне его равна нулю.

Вероятностные свойства случайных векторов ξ_1, \dots, ξ_N зависят только от разбиения параллелепипеда Π на элементарные, поэтому их генерацию можно осуществить, используя стандартный генератор случайных чисел.

Далее, из векторов ξ_i формируются векторы $\hat{\mathbf{x}}$, ПРВ которых является аппроксимирующая функция $\hat{w}(\mathbf{x})$.

3. Экспериментальные исследования

Расчет характеристик генератора (вычисление аппроксимации заданной плотности) является ресурсоемкой процедурой, особенно ситуация ухудшается в многомерном случае. Такие вычисления могут потребовать существенных вычислительных ресурсов, но их нужно выполнить один раз (для заданной ПРВ), в результате получается генератор с *гарантированной оценкой погрешности*, что выгодно отличает его от существующих методов, основанных на методе Метрополиса–Гастингса, такой оценкой не обладающих.

Учитывая это обстоятельство, проведенные экспериментальные исследования имеют скорее иллюстративный характер и направлены на подтверждение работоспособности и реализуемости предложенного метода, и демонстрации его преимуществ.

В рамках исследования использовались двумерные плотности усеченного нормального распределения, распределения Лагерра и Чебышева, реализация была проведена на платформе MATLAB. Для сравнения с алгоритмом Метрополиса–Гастингса использовалась его реализация на этой же платформе. Сравнение методов

проводилось с помощью вычисления относительной ошибки между теоретической W_i и эмпирической вероятностью F_i на сетке по формуле:

$$\delta = \frac{\sqrt{\sum_{i=1}^N (F_i - W_i)^2}}{\sqrt{\sum_{i=1}^N W_i^2}}.$$

Результаты исследований показали работоспособность предлагаемого метода и его превосходство по сравнению с методом Метрополиса–Гастингса по метрике (3).

Литература

1. J. von Neumann, Various techniques used in connection with random digits, in: Applied Mathematics Series / National Bureau of Standards, Vol. 12, United States Government Printing Office, Washington, DC, USA, 1951, pp. 36–38.
2. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, Equation of state calculations by fast computing machines, The journal of chemical physics 21 (6) (1953) 1087–1092.
3. R. Y. Rubinstein, D. P. Kroese, Simulation and the Monte Carlo method, Vol. 707, John Wiley & Sons, 2007.
4. S. Chib, E. Greenberg, Understanding the metropolis-hastings algorithm, The american statistician 49 (4) (1995) 327–335.
5. S. Chib, I. Jeliazkov, Marginal likelihood from the metropolis-hastings output, Journal of the American Statistical Association 96 (453) (2001) 270–281.
6. J. Geweke, H. Tanizaki, Note on the sampling distribution for the metropolis-hastings algorithm, Communications in Statistics-Theory and Methods 32 (4) (2003) 775–789.
7. D. D. Minh, D. L. Minh, Understanding the hastings algorithm, Communications in Statistics-Simulation and Computation 44 (2) (2015) 332–349.

UDC 004.9

Method of Random Vectors Generation with Given Probabilistic Characteristics

A. Y. Popkov*[†]

* *Institute for Systems Analysis of FRC CSC of RAS
44-2 Vavilova str., Moscow, 119333, Russian Federation*

[†] *Department of Information technologies
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: apopkov@isa.ru

It is proposed a method of random vectors generation with given probability density function. The method has guaranteed estimates of quality of generated random vectors. Experimental research of the method and its comparison with Metropolis-Hastings algorithm was conducted.

Key words and phrases: computer science, information technologies, conference proceedings.

УДК 004.4

Моделирование ценообразования на рынке смартфонов

А. Н. Пылаева

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
Ул. Миклуто-Маклая, д. 6, Москва, Россия, 141980*

Email: pylaevaalyona@gmail.com

Работа посвящена анализу зависимости формирования цены современного смартфона от ряда характеристик. На основе данных о двух ста пятидесяти смартфонах был проведен визуальный анализ, позволяющий выделить наиболее значимые переменные, а часть переменных исключить из анализа. Посредством корреляционного анализа были выявлены наиболее значимые по отношению к цене переменные. Основным инструментом являлся регрессионный анализ, посредством которого была построена регрессионная модель. В результате ряда исследований было выявлено, что построенная модель является качественной и адекватно описывает формирование цены современного смартфона в условиях рыночной экономики.

Ключевые слова: максимальная частота ядра, емкость аккумулятора, мультиколлинеарность факторов, автокорреляция в остатках, выбросы в данных.

1. Введение

Современный человек не мыслит свою жизнь без смартфона. Забыв свой «гаджет» дома, он чувствует себя словно без рук. Рынок же всячески подстраивается под современного человека: все быстрее и быстрее выходят новые смартфоны, все более функциональными они становятся, каждый производитель пытается изобрести что-то новое, полезное, неповторимое.

Целью работы является выявление зависимости между ценой смартфона и набором характеристик: известностью бренда, годом выпуска модели, объемом оперативной и внутренней памяти смартфона, количеством ядер и максимальной частотой ядра, емкостью аккумулятора, поддержкой сети LTE, диагональю, числом сим-карт, разрешением основной и фронтальной камер, а также наличием двойной камеры.

2. Основная часть

Первым шагом проведем регрессионный анализ. Построим модель линейной регрессии с имеющимися фиктивными (качественными) и количественными переменными. Коэффициенты первой модели представлены на рис. 1.

Как можно заметить, значимыми получились лишь 4 переменные: известность бренда, объем внутренней памяти, максимальная частота ядра и число сим-карт.

Построим вторую модель линейной регрессии, оставив в модели лишь значимые переменные, получим результаты, представленные на рис. 2 и 3.

Скорректированный коэффициент детерминации для второй модели получился равным 0,804. Все коэффициенты модели значимы на уровне значимости 5%.

Следующим шагом оценим качество выбранной регрессионной модели: проверим остатки модели на нормальность и гетероскедастичность.

Для проверки выборки на нормальность остатков перейдем к визуальной оценке: изучению построенной гистограммы плотности вероятности значений цены смартфона, а также исследованию графика квантилей:

Из анализа приведенных графиков можно сделать вывод, что распределение остатков модели не соответствует нормальному распределению.

Проверим модель для гетероскедастичности. Обозначим проверяемые гипотезы:

– H_0 — остатки гомоскедастичны;

Коэффициенты^а

Модель	Нестандартизованные коэффициенты		Стандартизованные коэффициенты	t	Значимость
	B	Стандартная ошибка	Бета		
1 (Константа)	-693,877	5728,841		-.121	.904
APPLE/SAMSUNG	10721,917	1559,037	.219	6,877	.000
RELEASE YEAR	448,406	612,706	.024	.732	.465
RAM_MB	-.178	.516	-.020	-.344	.731
MEMORY	.247	.016	.662	15,285	.000
NUMBER OF CORES	-194,884	247,080	-.030	-.789	.431
FREQUENCY_GGTS	7005,156	1393,066	.241	5,029	.000
BATTERY_Mah	.095	.354	.009	.270	.788
LTE	-246,371	1177,762	-.007	-.209	.834
DIAGONAL	324,269	918,689	.012	.353	.724
NUMBER OF SIM CARDS	-3753,049	2010,546	-.065	-1,867	.063
MAIN CAMERA_MP	-77,109	139,612	-.025	-.552	.581
FRONT-CAMERA_MP	45,961	147,162	.013	.312	.755
DUAL-CAMERA	1264,915	957,416	.042	1,321	.188

а. Зависимая переменная: COST (y), руб.

Рис. 1. Коэффициенты первой модели

Сводка для модели^б

Модель	R	R-квадрат	Скорректированный R-квадрат	Стандартная ошибка оценки	Дарбин-Уотсон
1	.898 ^а	.807	.804	5898,63770	1,826

а. Предикторы: (константа), NUMBER OF SIM CARDS, FREQUENCY, GGTS, APPLE/SAMSUNG, MEMORY

б. Зависимая переменная: COST (y), руб.

Рис. 2. Сводка для второй модели

Коэффициенты^а

Модель	Нестандартизованные коэффициенты		Стандартизованные коэффициенты	t	Значимость	95.0% Доверительный интервал для B	
	B	Стандартная ошибка	Бета			Нижняя граница	Верняя граница
1 (Константа)	1741,573	3974,986		.438	.662	-6087,934	9571,080
APPLE/SAMSUNG	10614,559	1489,396	.217	7,127	.000	7680,904	13548,215
MEMORY	.243	.014	.650	17,529	.000	.216	.270
FREQUENCY_GGTS	6440,980	1072,751	.222	6,004	.000	4327,988	8553,972
NUMBER OF SIM CARDS	-4098,900	1792,984	-.071	-2,286	.023	-7630,529	-567,271

а. Зависимая переменная: COST (y), руб.

Рис. 3. Коэффициенты второй модели

– H1 — остатки гетероскедастичны.

Для проверки модели на гетероскедастичность воспользуемся статистическим пакетом Eviews. Результаты проверки представлены на рис. 6:

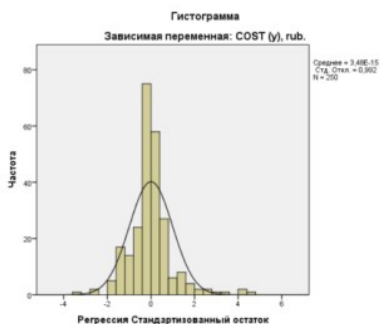


Рис. 4. Гистограмма

Нормальный P-P график регрессии Стандартизованный остаток

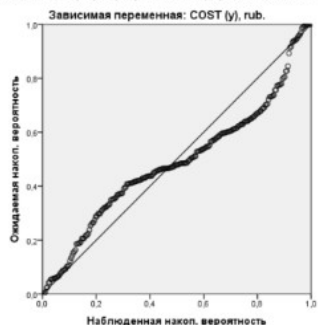


Рис. 5. График квантилей

Heteroskedasticity Test: White

F-statistic	18.32659	Prob. F(12,237)	0.0000
Obs*R-squared	120.3272	Prob. Chi-Square(12)	0.0000
Scaled explained SS	383.8331	Prob. Chi-Square(12)	0.0000

Test Equation:

Dependent Variable: RESID*2

Method: Least Squares

Date: 12/18/17 Time: 22:01

Sample: 1 250

Included observations: 250

Collinear test regressors dropped from specification

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.68E+08	1.53E+08	1.756920	0.0802
APPLE_SAMSUNG*2	-3.52E+08	1.23E+08	-2.855634	0.0047
APPLE_SAMSUNG*MEMORY	-1582.383	706.7900	-2.238830	0.0261
APPLE_SAMSUNG*FREQUENCY	2.77E+08	49557938	5.583441	0.0000
APPLE_SAMSUNG*NUMBER_OF_SIM...	10028748	45425028	0.220776	0.8255
MEMORY*2	0.002869	0.002371	1.210260	0.2274
MEMORY*FREQUENCY	324.4734	548.7098	0.591339	0.5549
MEMORY*NUMBER_OF_SIM_CARDS	227.4361	650.0200	0.349891	0.7267
MEMORY	-186.1778	1659.817	-0.112168	0.9108
FREQUENCY*2	69909793	27715294	2.522427	0.0123
FREQUENCY*NUMBER_OF_SIM_CAR...	31085723	55143162	0.563728	0.5735
FREQUENCY	-3.15E+08	1.44E+08	-2.187809	0.0297
NUMBER_OF_SIM_CARDS*2	-16938912	33438564	-0.506568	0.6129

Рис. 6. Реализация теста Уайта

Результаты показали, что наблюдается гетероскедастичность остатков.

3. Заключение

Обобщая всё выше сказанное, можно сделать следующие выводы:

- остатки полученной регрессионной модели не являются нормальными;
- в модели присутствует гетероскедастичность, однако подправка показала, что на значимость коэффициентов она не повлияла.

Если говорить об интерпретации коэффициентов, то можно заметить, что мы получили достоверные результаты:

- Цена смартфона известного бренда (Samsung, Apple) на 10 614 рублей 56 копеек выше, чем цена смартфона неизвестного производителя с аналогичными характеристиками.
- Объем встроенной памяти смартфона не оказывает значимого влияния на цену (в совокупности с другими факторами).
- Максимальная частота ядра, как видим, оказывает значимое влияние на цену. Так, при увеличении частоты на 1ГГц цена смартфона увеличивается на 6 440 рублей 98 копеек.
- Если говорить о такой характеристике, как наличие слота для двух сим-карт, то мы также видим достаточно сильное влияние на величину цены. Так смартфон, имеющий слот для двух сим-карт, стоит на 4 098 рублей 90 копеек дешевле смартфона с одной сим-картой.

Литература

1. Нэреш К. Малхорта Маркетинговые исследования. Маркетинговые исследования. Практическое руководство, 3-е издание.: Пер. с англ. — М.: Издательский дом «Вильямс», 2002. — 960 с.
2. Моосмюллер Г., Ребик Н. Н. Маркетинговые исследования с SPSS.: Учеб. пособие. — М.: ИНФРА-М, 2009. — 160 с. — (Высшее образование).
3. Беляевский И. К. Маркетинговое исследование: Учебное пособие, руководство по изучению дисциплины, практикум по курсу, учебная программа / Московский государственный университет экономики, статистики и информатики. — М. 2004. — 414 с.

UDC 004.4

Modeling of Price Formation in the Smartphone Market

A. N. Pylaeva

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: pylaevaalyona@gmail.com

The work is devoted to the analysis of the dependence of price of a modern smartphone on a number of characteristics. Based on data of two hundred and fifty smartphones, a visual analysis was carried out, which allows selecting the most significant variables, and excluding some of the variables from the analysis. By means of the correlation analysis, the most significant variables relative to the price were identified. The main instrument was regression analysis, through which a regression model was built. As a result of studies, it was revealed that the model constructed is qualitative and adequately describes the formation of the price of a modern smartphone in a market economy.

Key words and phrases: maximum core frequency, battery capacity, multicollinear factors, autocorrelation in residuals, emissions in data.

УДК 004.4

Анализ цен на ноутбуки с помощью эконометрических моделей

Д. А. Пяткина, С. И. Матюшенко, П. А. Зотова, А. П. Казандаева

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: pyatkina_da@rudn.university, matyushenko_si@rudn.university, zotova-pa@mail.ru, anna@kazandaeva.ru

В работе рассматривается алгоритм построения эконометрических моделей по данным для цен на ноутбуки и их характеристикам. Подробно рассмотрена методология такого исследования, приведён ряд графиков и выводов. Для исследования применён эконометрический пакет Eviews. Важным моментом анализа таких моделей является проверка остатков моделей на нормальность и гетероскедастичность, а также разбиение выборки на подвыборки для получения более адекватных результатов. Основной трудностью при построении таких моделей является момент сбора данных и кодирования переменных для дальнейшей работы. Очень важен предварительный визуальный анализ, который часто позволяет увидеть некоторые закономерности ещё до численных расчётов.

Ключевые слова: эконометрическая модель, нормальность остатков, прогноз, фиктивные переменные, значимость, подправка.

1. Введение

Рассматриваются данные по 193 ноутбукам, собранные по различным торговым сетям с целью анализа цен на ноутбуки в зависимости от их характеристик. Описание переменных (большинство из которых – фиктивные):

- PROCESSOR = 1 для AMD и 0 для остальных ноутбуков
- IGROV = 1 для игровых компьютеров
- DIAG – размер диагонали
- VEC – вес ноутбука
- OPER_SYST кодируется в зависимости от опер системы 4 значениями
- Материал корпуса кодируется 4 значениями
- DISK – объём жёсткого диска

2. Основная часть

Построена и отобрана следующая модель регрессии [1].

Интерпретация коэффициентов полулогарифмической модели и анализ качества модели :

- При увеличении диагонали на 1% цена увеличивается в среднем на 1,42%
- При увеличении веса на 1 кг цена в среднем уменьшается на 27 тыс. рублей
- Ноутбуки с полностью металлическим корпусом стоят в среднем на 62 тыс. дороже ноутбуков с пластиковым корпусом.
- Ноутбуки с операционной системой Windows стоят в среднем на 81 тыс. дешевле, чем МАК-буки.
- Игровые компьютеры в среднем стоят на 56 тыс. дороже неигровых
- Цена с ростом объёма жесткого диска уменьшается – видимо это связано с тем, что есть компьютеры вообще без жесткого диска.

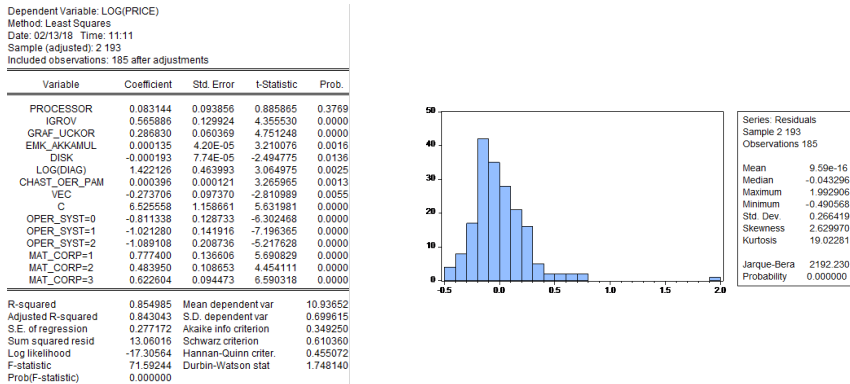


Рис. 1. Логарифмическая модель для всей совокупности ноутбуков
+ проверка на нормальность

Skewness=2,62 – модель занижает цену.

В дальнейшем предполагается провести исследование зависимости цены от фирмы производителя. Анализ остатков (нормальности остатков нет, но некоторая симметричность присутствует)

Пересчёт стандартных ошибок – корректировка по Уайту. Все переменные, кроме процессор остаются значимыми. Значимость переменной процессор уменьшилась. Модель значима в целом на 99%, DW – высокий, что свидетельствует об отсутствии автокорреляции. Качество модели по MAPE – среднее: ошибка 17%. Ниже приведён график остатков модели – видно, что модель плохо прогнозирует дорогой сегмент.

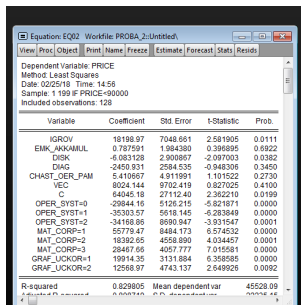


Рис. 2. Модель для дешёвого сегмента (отбор наблюдений)

Также наши данные позволяют провести исследование по отдельным фиктивным переменным. На рисунке ниже приведён разброс цен для ноутбуков в зависимости от материала корпуса. Видно, что в модели есть выбросы. Дальнейший путь исследования – это чистка от выбросов.

3. Заключение

Если построить модель для данных, в которых цена < 90000 рублей – модель становится более адекватной и объясняемой [2]. Здесь мы берём линейную модель.

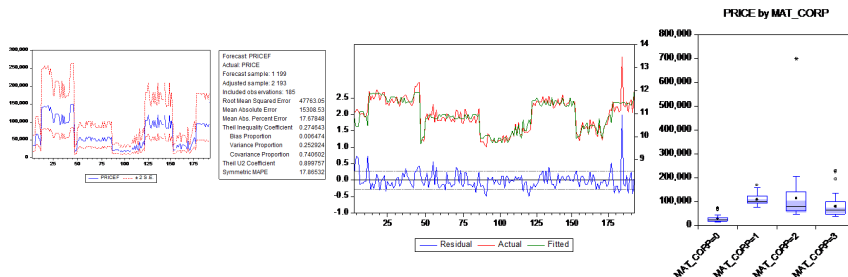


Рис. 3. Графики остатков, прогнозных значений и исходных данных, зависимость цены от материала корпуса

Выводы: для игрового ноутбука цена в среднем на 18 тыс. выше, чем обычного, при увеличении диагонали на 1 дюйм цена в среднем уменьшается на 2450 руб., при увеличении частоты операционной памяти на 1 ед. цена увеличивается в среднем на 5 тыс. рублей, также цена ноутбука зависит от операционной системы, металлический корпус вызывает рост цены в среднем на 28 тысяч по сравнению с пластмассовым. У этой модели лучше характеристики по качеству остатков и по MAPE=16%.

Литература

1. Магнус Я.Р. Эконометрика. Начальный курс: Учебное пособие для вузов / Я. Р. Магнус, П. К. Кагышев, А. А. Пересецкий. — 2-е изд., испр. — М.: Дело, 2010.
2. Терещенко О.В., Курилович Н.В., Князева Е.И. Многомерный статистический анализ данных в социальных науках. — Минск, БГУ, 2012.

UDC 004.4

Analysis of Notebook Prices with Econometric Models

D. A. Pyatkina, S. I. Matyushenko, P. A. Zotova, A. P. Kazandaeva

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: pyatkina_da@rudn.university, matyushenko_si@rudn.university, zotova-pa@mail.ru, anna@kazandaeva.ru

The paper considers an algorithm for constructing econometric models based on data for notebook prices and their characteristics. The methodology of such research is considered in detail, a number of graphs and conclusions are presented. For the study, the econometric package Eviews was applied. An important point in the analysis of such models is the verification of model residues for normality and heteroscedasticity, as well as splitting the sample into subsamples to obtain more adequate results. The main difficulty in constructing such models is the time of data collection and coding of variables for further work. Very important is the preliminary visual analysis, which often allows you to see some patterns even before numerical calculations.

Key words and phrases: econometric model, residual normality, forecast, fictitious variables, significance, correction.

УДК 519.246.5

Оценка зависимостей тяжелых хвостов при помощи аппарата копул для анализа осадков

Н. Д. Рассахан*, Е. Ю. Щетинин†

* Кафедра прикладной математики

Московский государственный технологический университет «Станкин»
Вадковский пер., д. 3а, Москва, Россия, 127055

† Всероссийский научно-исследовательский институт
по проблемам гражданской обороны и чрезвычайных ситуаций МЧС России
ул. Давыдовская, д. 7, Москва, Россия, 121352

Email: rassahan@gmail.com, riviera-molto@mail.ru

Измерение хвостовой зависимости является важной задачей во многих прикладных науках для оценивания риска совместного наступления экстремальных событий. Обычно мерой зависимости является коэффициент хвостовой зависимости. Корреляция Пирсона не является подходящей мерой для оценивания зависимости двух величин в контексте совместного наступления экстремальных событий, когда они представляют интерес для исследователя, так как она учитывает экстремальные события так же (с тем же весом), что и «рядовые» события, хотя зависимость между экстремальными событиями может сильно отличаться от общей картины.

Данная работа подчеркивает важность учитывания хвостовой зависимости в контексте двумерного анализа при помощи копул. В связи с учащающимися природными катаклизмами резко встает вопрос об оценивании разнотипных рисков (в т.ч. экономических) и последствий их совместного наступления с учетом пространственных связей между наблюдениями. Сравниваются 2 непараметрические оценки коэффициента хвостовой зависимости для оценки зависимости между ежедневными наблюдениями осадков в городах Европейской части России. Большинство существующих оценок зависит от порога k и, следовательно, при выборе используемого значения k происходит трейдофф между смещением и вариацией. Для установления баланса в работе представлен алгоритм, основанный на использовании скользящего среднего и поиска «стабильного участка» коэффициента хвостовой зависимости. Именно среднее значение на «стабильном» участке и принимается за значения оценки коэффициента хвостовой зависимости.

Ключевые слова: теория экстремальных величин, пространственное моделирование, экстремальные осадки, пространственные структуры статистической зависимости, коэффициент хвостовой зависимости.

1. Введение

Важнейшей частью анализа многомерных экстремальных величин является исследование их экстремальных зависимостей, в основном использующее для этого коэффициенты хвостовой зависимости. Для двумерного вектора (X_1, X_2) коэффициент верхней хвостовой зависимости имеет вид

$$\lambda_U = \lim P(F_1(X_1) > t | F_2(X_2) > t), \quad t \rightarrow 1^-, \quad (1)$$

где F_1, F_2 — частные функции распределения случайных величин X_1, X_2 соответственно, $0 < t \leq 1$ — некоторый порог. Можно также определить для двумерного вектора коэффициент нижней хвостовой зависимости

$$\lambda_L = \lim P(F_1(X_1) < t | F_2(X_2) < t), \quad t \rightarrow 0^+. \quad (2)$$

С использованием функции копулы [3] выражения (1), (2) принимают следующий вид:

$$\lambda_U = \lim_{t \rightarrow 1^-} \frac{1 - 2t + C(t, t)}{1 - t}; \quad \lambda_L = \lim_{t \rightarrow 0^+} \frac{C(t, t)}{t}. \quad (3)$$

2. Основная часть

Методы оценивания коэффициента хвостовой зависимости являются важными и необходимыми инструментами анализа структур экстремальных осадков, исследуемых в данной работе. Опишем некоторые из них. Прежде всего, таковыми являются непараметрические оценки, в основе которых использована концепция эмпирической копулы $C^{(n)}(u, v)$.

Пусть $(X_1^{(1)}, X_2^{(1)}), \dots, (X_1^{(n)}, X_2^{(n)})$ - независимые одинаково распределенные копии двумерного случайного вектора (X_1, X_2) . Используя

$$\lambda = 2 - \lim_{t \rightarrow 1^-} \frac{1 - C(t, t)}{1 - t} = 2 - \lim_{t \rightarrow 0^+} \frac{1 - C(1 - t, 1 - t)}{t}, \quad (4)$$

получим следующую оценку коэффициента верхней хвостовой зависимости (1):

$$\hat{\lambda}^{SEC} \equiv \hat{\lambda}^{SEC}(k) = 2 - \frac{1 - \hat{C}\left(1 - \frac{k}{n}, 1 - \frac{k}{n}\right)}{\frac{k}{n}}, \quad 1 \leq k < n. \quad (5)$$

Учитывая $\log(1 - t) \sim -t$ при $t \approx 0$, получим другую оценку:

$$\hat{\lambda}^{LOG} \equiv \hat{\lambda}^{LOG}(k) = 2 - \frac{\log \hat{C}\left(1 - \frac{k}{n}, 1 - \frac{k}{n}\right)}{\log\left(1 - \frac{k}{n}\right)}, \quad 1 \leq k < n.$$

Здесь \hat{C} представляет собой эмпирическую копулу, определенную как

$$\hat{C}\left(1 - \frac{k}{n}, 1 - \frac{k}{n}\right) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}\{\hat{F}_1(X_1^{(i)}) \leq 1 - \frac{k}{n}, \hat{F}_2(X_2^{(i)}) \leq 1 - \frac{k}{n}\}, \quad 1 \leq k < n$$

где $\mathbb{1}$ является индикаторной (характеристической) функцией, а $\hat{F}_j, j = 1, 2$ являются маргинальными эмпирическими функциями распределений X_1 и X_2 соответственно.

Как видно из выражений (4), (5), оценки зависят от порога k , выбор которого определяется, согласно теореме о стабильности оценки λ_U [2], путем выбора достижения баланса между значениями ее смещения и вариации. При достаточно большом размере данных n процедуру установления баланса оценки λ_U характеризует участок графика, на котором она постоянна. Для нахождения такого участка в работе [1] предложен алгоритм, состоящий в следующем:

1. Эмпирическая оценка сглаживается с использованием скользящего среднего с окном шириной $b = \text{int}(0.05n)$, в результате чего получается последовательность $\hat{\lambda}_1, \dots, \hat{\lambda}_{n-2b}$.
2. Из последовательности $\hat{\lambda}_1, \dots, \hat{\lambda}_{n-2b}$ последовательным перебором выделяется вектор, состоящий из величин $(\hat{\lambda}_k, \dots, \hat{\lambda}_{k+m-1})$, где $k = 1, \dots, n - 2b + m - 1, m = \text{int}(\sqrt{n - 2b})$
3. Если текущий вектор удовлетворяет условию $\sum_{i=k+1}^{k+m-1} |\bar{\lambda}_i - \bar{\lambda}_k| \leq 2\sigma$, где σ - среднеквадратичное отклонение вектора $\hat{\lambda}_1, \dots, \hat{\lambda}_{n-2b}$, то итоговая оценка имеет вид $\lambda_U = \frac{1}{m} \sum_{i=1}^m \bar{\lambda}_{k+i-1}$. Если условие не выполнено после перебора, то $\lambda_U = 0$.

Для реализации графика выбрана пара городов Можайск–Москва, полученное значение $\hat{\lambda}_U = 0,7824$, что свидетельствует о наличии зависимости между экстремальными событиями для данной пары городов.

3. Заключение

Из текста данной работы становится ясной необходимость учитывать коэффициент хвостовой зависимости при анализе многомерных экстремальных величин с использованием копул. Были рассмотрены и сравнены две непараметрические оценки коэффициента верхней хвостовой зависимости $\hat{\lambda}_U^{SEC}$, $\hat{\lambda}_U^{LOG}$.

Рассмотренная теоретическая база применена к ежедневным данным по осадкам в Европейской части России, в частности, в Можайске и Москве. Для данных городов при помощи алгоритма получены оценки коэффициента верхней хвостовой зависимости на примере выражения (3).

Литература

1. Cole S., Heffernan J., Tawn J. Dependence Measures for Extreme Value Analysis // *Extremes*. — 1999. — Vol. 2 (4). — Pp. 339–365.
2. Genest C., Ghoudi K., Rivest L. P. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions // *Biometrika*. — 1995. — Vol. 82 (3). — Pp. 543–552.
3. Grimaldi S., Serinaldi F. Asymmetric copula in multivariate flood frequency analysis // *Adv. Water Resour.* — 2006. — Vol. 29 (8). — Pp. 1155–1167.

UDC 519.246.5

Estimation of Heavy Tail Dependence Based on Copulas for the Precipitation Analysis

N. D. Rassakhan*, E. Yu. Shchetinin[†]

* *Department of Applied Mathematics
Moscow State Technology University "STANKIN"
3a Vadkovsky Ln., Moscow, 127055, Russian Federation*

[†] *FGU "All-Russian research institute on problems of civil defence and emergencies"
of Emergency Control Ministry of Russia
7 Davydkovskaya St., Moscow, 121352, Russian Federation*

Email: rassahan@gmail.com, riviera-molto@mail.ru

Consideration of tail dependence is a very important part of risk analysis in many applied sciences that is measured in order to estimate the risk of simultaneous extreme events. Usually the tail dependence coefficient is the measurement in question. Pearson correlation coefficient unfortunately is not a suitable measure for estimating dependencies between two quantities in the context of simultaneous occurrence of extreme events when these events are of interest for the researcher because it takes extreme events into account with the same weight as it takes "normal" events although extreme events' dependence may slightly differ.

Present work emphasizes the importance of taking into consideration tail dependencies in bivariate statistical analysis using copulas. Due to increasing frequency of environmental cataclysms the issue of analyzing risks (e.g. economic losses) and their consequences comes to the fore. Moreover, researchers should take into account consequences of their joint occurrence. Two non-parametric estimators of tail dependence coefficients were compared in order to estimate correlation between daily cumulative rainfall totals recorded in central European part of Russia. Major part of existing estimators depends on threshold k and thus there is a trade-off between variance and bias during the calculation of the best value for k . For balancing an algorithm is presented that is based on using moving average and then searching the "stable" part of tail dependence coefficient. Estimate of tail dependence coefficient is assumed to be equal to mean value on the "stable" part.

Key words and phrases: extreme value theory, spatial modelling, extreme precipitation, spatial structures of statistical dependence, tail dependence coefficient.

UDC 574.34

Analysis of Stochastics Dynamic Population Growth

L. A. Sevastianov, D. G. Vasilyeva

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: sevastianov_la@rudn.university, vassilyeva.darya@yandex.ru

The random walk and drift model of population dynamics is an important approach in conservation biology, partly because its parameters are easily estimated from periodic observations of population size. Estimating the model with noisy data is problematic, however, because the commonly used estimators of process variation are biased if population abundance measurements are imprecise, and a recently developed method that attempts to remove this bias is not robust. The likelihood function allows the variances of the process error and measurement error and the growth rate of the population to be estimated in a way that is robust and fully supported by statistical theory. Comparative analysis using simulated data indicates that the Kalman-filter method reduces the bias in estimates of process variance without yielding negative variance estimates. Demographic noise causes unlimited population growth in a broad class of models which, without noise, would predict a stable finite population. We study this effect on the example of a stochastic birth-death model which includes immigration, binary reproduction and death. The unlimited population growth is considered using the computer simulation methods.

Key words and phrases: stochastics dynamic, population growth, computer simulation.

1. Introduction

Since the celebrated essay of Malthus [3] quantitative modeling of population dynamics has attracted much interest. To a large extent, this interest is powered by the danger of a Malthusian catastrophe, when too a rapid population growth causes a fatal lack of resources. Here we focus on a variant of Malthusian catastrophe by considering not too a large population that undergoes binary reproduction, immigration and death. Although macroscopically stable, this population can be pushed to the Malthusian limit (a critical population size that sparks a Malthusian catastrophe) by rare large fluctuations. We show that unlimited population growth proceeds as a slow decay of a metastable probability distribution (MPD) of the population. We determine the MPD and the decay time analytically by developing a systematic WKB theory for the master equation and combining it with the van Kampen system size expansion. We show that the MPD is described by two different WKB modes which are strongly coupled in a narrow region around the unstable fixed point of the deterministic rate equation of the model. At large population sizes the MPD exhibits a power-law tail so that all the distribution moments, except the zeroth one, diverge.

At the deterministic level of modeling a Malthusian catastrophe does not occur if the gain and loss processes balance each other so that the resulting steady-state population size is stable with respect to small perturbations. Real populations, however, behave stochastically, rather than deterministically [1]. The stochasticity may cause an unlimited population growth in a broad class of models which deterministic counterparts predict a stable finite population size [5]. We will investigate this remarkable phenomenon on the example of a birth-death model [2] which accounts for binary reproduction with rate λ , immigration with rate σ , and death with rate μ .

The rate equation for this model is

$$\dot{n} = \sigma - \mu n + 0.5 \lambda n^2, \quad (1)$$

where $n(t) \gg 1$ is the average population size.

We develop an approach that has been proposed by B. Meerson, P. V. Sasorov [4], We study this effect on the example of a stochastic birth-death model which includes immigration, binary reproduction and death. The unlimited population growth is considered using the computer simulation methods.

2. Analysis of stochastic dynamic population growth

For a relatively low death rate [4], $\mu^2 < 2\sigma\lambda$, (1) does not have fixed points, and the population size blows up in finite time for any $n(t=0)$. For $2 > 2\sigma\mu$ (1) has two fixed points: $n_1 = \Omega(1 - \delta)$ and $n_2 = \Omega(1 + \delta)$, where $\Omega = \mu/\lambda \gg 1$ and $\delta^2 = 1 - 2\sigma\lambda/\mu^2$. When starting from any $n(t=0) < n_2$, the population size flows to the attracting fixed point $n = n_1$ with a characteristic relaxation time $\tau_r = 1/(\mu\delta)$, and stays there forever. Master equation, absorbing state and decay of metastable state. The demographic noise, ignored by the rate equation (1), is accounted for by the master equation, see [5], which governs the evolution of probability $P_n(t)$ to have n individuals at time t :

$$\dot{P}_n = \lambda_{n+1}P_{n-1} - (\lambda_n + \mu_n)P_n + \mu_{n+1}P_{n+1}, \quad (2)$$

where $\lambda_n = (\lambda/2)n(n-1) + \sigma$, and $\mu_n = \mu n$. One striking property of (2) concerns its steady state.

Summing up the first n equations in (2), we obtain

$$\sum_{l=0}^n \dot{P}_l(t) = -\frac{n(n-1)}{2\Omega}P_n + (n+1)P_{n+1} - \frac{\gamma\Omega}{2}P_n,$$

where $\gamma = 1 - \delta^2 = 2\sigma\lambda/\mu^2$, and the time is rescaled by the death rate: $\mu t \rightarrow t$. Putting $d/dt P_l(t) = 0$, we obtain

$$P_{n+1} = -\frac{n(n-1) + \gamma\Omega^2}{2\Omega(n+1)}P_n.$$

Clearly, $\lim_{n \rightarrow \infty} P_n = \infty$ unless $P_n = 0$, $n = 0, 1, 2, \dots$. Therefore, the ultimate state of the stochastic process corresponds to an empty system, in a stark contrast to the prediction of the rate equation (1).

We study this effect on the example of a stochastic birth-death model (1) which includes immigration, binary reproduction and death using the computer simulation methods (see Fig. 1).

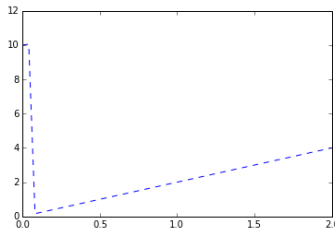


Figure 1. Stochastic birth-death model computer simulation

The unlimited population growth for birth-death model (1) is considered by using the computer simulation methods and by using a simple birth-death process as an

example, we have using a systematic theory of unlimited population growth driven by demographic growth.

3. Conclusions

The unlimited population growth is considered by using the computer simulation methods and by using a simple birth-death process as an example, we have developed a systematic theory of unlimited population growth driven by demographic growth.

Acknowledgments

The publication has been prepared with the support of the "RUDN University Program 5-100".

References

1. M. S. Bartlett, Stochastic Population Models in Ecology and Epidemiology, Wiley, New York, 1961.
2. V. Elgart and A. Kamenev, Phys. Rev. E 70, 41106, 2004.
3. T. R. Malthus, An Essay on the Principle of Population, Murray, London, 1798.
4. B. Meerson, P. V. Sasorov, Noise driven unlimited population growth, Phys. Rev. E 78, 060103(R), 2008.
5. N. G. van Kampen, Stochastic Processes in Physics and Chemistry, North-Holland, Amsterdam, 2001.

УДК 574.34

Стохастическая динамика роста популяций

Л. А. Севастьянов, Д. Г. Васильева

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклухо-Маклая, д. 6, Москва, Россия, 117198*

Email: sevastianov_la@rudn.university, vasilyeva.darya@yandex.ru

Модель динамики популяции с использованием случайного блуждания с дрейфом является важным инструментом в биологии по причине того, что параметры модели легко оцениваются по периодическим наблюдения за численностью популяций. Однако анализ параметров моделей популяционной динамики с зашумленными данными не всегда возможен, поскольку используемые оценки случайных процессов в биологии не всегда достаточно неточны, а методы, который пытается устранить эти недостатки, не является достаточно надежными. Использование методов функции правдоподобия позволяет оценивать отклонения в динамике роста популяции, оценивать погрешности измерения численности особей и темпы роста популяции с более высокой точностью. Сравнительный анализ с использованием методов компьютерного моделирования показывает, что использование методов фильтрации с применением фильтра Кальмана уменьшает смещение в оценках дисперсии процессов и исключает получения отрицательных оценок дисперсии. Наличие стохастической компоненты в параметрах модели популяционной динамики в виде шума вызывает неограниченный рост популяции в широком классе моделей, которые без такого шума давали решения в виде ограниченной динамики популяции. Мы исследуем стохастическую динамику роста популяций на примере модели рождения и смерти, которая включает в себя иммиграцию, рождение и смерть особей. С помощью компьютерного моделирования исследована стохастическая динамика роста популяций.

Ключевые слова: стохастическая динамика, рост популяции, компьютерное моделирование.

УДК 517.977;681.5

Оценка состояния нелинейных сетевых систем управления, представленных нечеткими моделями Такаги–Сугено

Ю. В. Талагаев

*Кафедра физики и информационных технологий
Балашовский институт (филиал)*

*ФГБОУ ВО «Саратовский национальный исследовательский государственный университет имени Н. Г. Чернышевского»
ул. Карла Маркса, д. 29, Балашов, Россия, 412309*

Email: ytalagaev@yandex.ru

Рассмотрен подход к описанию нелинейных сетевых систем управления, основанный на их представлении в форме нечеткой модели Такаги–Сугено. Возможные неопределенности, сопровождающие функционирование системы, учтены включением в нечеткую модель внешних ограниченных возмущений. Показано, что задача оценки состояния системы может быть эффективно решена на основе техники инвариантных эллипсоидов.

Ключевые слова: нелинейная сетевая система управления, нечеткая модель Такаги–Сугено, оценка состояния, инвариантные эллипсоиды.

1. Введение

Развитие компьютерных и сетевых технологий влечет за собой эволюцию традиционных систем управления к современным сетевым системам управления, которые обладают большей гибкостью и надежностью, менее энергозатратны, просты в установке и обслуживании [1]. Принципиальная особенность нелинейной сетевой системы управления (НССУ) состоит в том (см. рис. 1), что она интегрирует информацию, сообщения и управление в единую систему в которой контуры управления замкнуты через коммуникационные сети (проводные, беспроводные, IP-сети и др.). Реализация НССУ на практике приводит к следующим трудностям: 1) вызванные сетью задержки влияют на точность зависящих от времени вычислений (оценка состояния и др.) и могут снижать производительность управления (искажение сигналов регулятора); 2) в случае перегруженности сетевого трафика и/или ограниченной надежности сети возможна потеря или задержка данных, что приводит к необходимости управления в условиях неполной информации.

В течение последнего десятилетия разработаны различные подходы к моделированию и управлению НССУ (delay-dependent, stochastic, switched control и др.), позволяющие учитывать сетевые задержки, потерю данных и ограничения пропускной способности каналов связи. Свою продуктивность показал подход, основанный на описании НССУ средствами нечеткой логики, которая уже давно зарекомендовала себя как альтернативное средство анализа устойчивости и управления сложными нелинейными системами [2,3]. В основе этого подхода лежит аппроксимация или эквивалентная замена в определенной области фазового пространства данной системы нечеткой моделью Такаги–Сугено (Т–С) [4], которая представляет собой набор линейных моделей, взвешенных функциями принадлежности. Линейность составляющих нечеткую Т–С модель подсистем делает возможным применение методов анализа и синтеза линейных систем, в том числе аппарата линейных матричных неравенств (ЛМН).

В настоящей работе исследуется ситуация, когда нелинейный объект сетевого управления описывается нечеткой Т–С моделью, подверженной действию внешних ограниченных возмущений. Включение в модель возмущений является одним из способов учета неопределенности, свойственных НССУ [5]. Важной задачей в таком случае является оценка состояния системы. В работе показывается, что эффективным способом ее решения является техника инвариантных эллипсоидов [6], позволяющая аппроксимировать инвариантное множество системы.

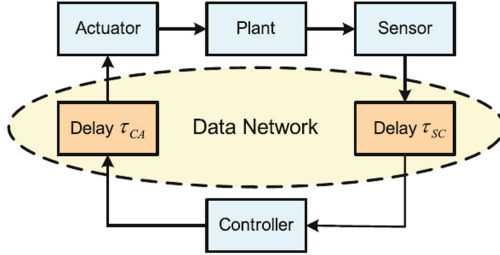


Рис. 1. Структура сетевой системы управления [1]

2. Основной результат

Пусть разомкнутая НССУ описывается следующей нечеткой Т–С моделью

Plant rule i :

IF $\theta_1(t)$ is M_{i1} and ... and $\theta_p(t)$ is M_{ip} ,

THEN $\dot{x}(t) = A_i x(t) + D_i w(t)$, $x(0) = x_0$,

где $x(t) \in R^n$ — состояние системы, $w(t) \in R^r$ — внешнее возмущение, удовлетворяющее ограничению $\|w(t)\| \leq 1$ для всех $t \geq 0$, $i = 1, \dots, q$ — номер нечеткого правила IF–THEN, $A_i \in R^{n \times n}$, $D_i \in R^{n \times r}$ — известные постоянные матрицы, $\theta_i(t), \dots, \theta_p(t)$ — переменные послыки, образующие вектор $\theta(t) = (\theta_1(t), \dots, \theta_p(t))^T$, M_{ik} — нечеткие множества ($k = 1, \dots, p$). При нечетком смешивании общая нечеткая Т–С модель выводится следующим образом [4]:

$$\dot{x}(t) = \sum_{i=1}^q h_i(\theta(t)) (A_i x(t) + D_i w(t)), \quad x(0) = x_0, \quad \|w(t)\| \leq 1, \quad (1)$$

где $h_i(\theta(t))$ — функции принадлежности ($0 \leq h_i(\theta(t)) \leq 1$, $\sum_{i=1}^q h_i(\theta(t)) = 1$).

Для действующего на нечеткую Т–С систему (2) возмущения $w(t)$ предполагается только то, что оно ограничено. Никаких других ограничений не накладывается. Это позволяет учесть в модели возможные неопределенности, сопровождающие функционирование НССУ. Система предполагается устойчивой, и задача состоит в оценке влияния возмущений на ее состояние. Естественным подходом к решению сформулированной задачи является обращение к концепции инвариантных множеств и варианту их аппроксимации — инвариантным эллипсоидам [6]. Выбор эллипсоидов удобен из-за их связи с общей квадратичной функцией Ляпунова нечеткой Т–С модели и возможностью использования аппарата ЛМН.

Определим для устойчивой нечеткой Т–С модели (2) с общей квадратичной функцией Ляпунова $V(x(t)) = x(t)^T P^{-1} x(t)$ семейство инвариантных эллипсоидов: общий для всех линейных подсистем в (2) эллипсоид с центром в начале координат $E_x = \{x \in R^n : x^T P^{-1} x \leq 1\}$, $P \succ 0$, называется *инвариантным* по состоянию для (2), если из $x_0 \in E_x$ следует $x(t) \in E_x$ для всех $t \geq 0$.

При оценке состояния системы особенно интересны минимальные эллипсоиды. В качестве критерия минимальности можно выбрать минимизацию суммы квадратов полуосей эллипсоида. Тогда матрица \hat{P} минимального инвариантного эллипсоида E_x нечеткой Т–С модели (2) при фиксированном α определяется решением задачи полуопределенного программирования

$$\min_{P=P^T, \alpha>0} tr(P) \quad (2)$$

с ограничениями

$$A_i P + P A_i^T + \alpha P + a^{-1} D_i D_i^T \leq 0, \quad \begin{pmatrix} 1 & x_0^T \\ x_0^T & P \end{pmatrix} \succeq 0, \quad P \succ 0, \quad i = 1, \dots, q,$$

В работе приведены результаты решения сформулированной задачи, реализованные в программном пакете CVX среды MATLAB, демонстрирующие эффективность представленного подхода для эталонных систем. Низкая степень консервативности оценок инвариантного множества подтверждена сравнением с предложенным в [7, 8] альтернативным методом оценки состояния нечеткой T-C модели.

Литература

1. Peng C., Yue D., Han Q.-L. Communication and Control for Networked Complex Systems. — Berlin: Springer, 2015.
2. Feng G. A survey on analysis and design of model-based fuzzy control systems, IEEE Trans. Fuzzy Syst. — 2006. — Vol. 14 (5). — Pp. 676–697. <http://dx.doi.org/10.1109/TFUZZ.2006.883415>
3. Дружинина О. В., Масина О. Н. О подходах к анализу устойчивости нелинейных динамических систем с логическими регуляторами, Современные информационные технологии и ИТ-образование. — 2017. — Vol. 13 (2). — С. 40–49.
4. Tanaka K., Wang H. O. Fuzzy Control Systems Design and Analysis: A Linear Matrix Inequality Approach. — New York: Wiley 2001.
5. Jiang X., Han Q.-L. On Designing Fuzzy Controllers for a Class of Nonlinear Networked Control Systems // IEEE Trans. Fuzzy Syst. — 2008. — 16 (4). — Pp. 1050–1060. <https://doi.org/10.1109/TFUZZ.2008.917293>
6. Хлебников М. В., Поляк Б. Т., Кунцевич В. М. Оптимизация линейных систем при ограниченных внешних возмущениях (техника инвариантных эллипсоидов) // Автоматика и телемеханика. — 2011. — № 11. — Pp. 9–59.
7. Талагаев Ю. В. Анализ и синтез сверхустойчивых нечетких систем Такаги–Сугено // Проблемы управления. — 2016. Vol. 6. — Pp. 2–11.
8. Talagaev Y. State Estimation and Stabilization of Continuous-Time Takagi–Sugeno Fuzzy Systems with Constraints of Positiveness and Superstability // IEEE Conference on Fuzzy Systems, 2017. <https://doi.org/10.1109/FUZZ-IEEE.2017.8015437>

UDC 517.977;681.5

State Estimation of the Nonlinear Network Control Systems Represented by Takagi-Sugeno Fuzzy Models

Y. V. Talagaev

*Department of Physics and Information Technologies
Balashov Institute of Saratov State University
29 Karl Marx str., Balashov, 412309, Russian Federation*

Email: ytalagaev@yandex.ru

An approach to the description of nonlinear network control systems based on their representation in the form of a Takagi-Sugeno fuzzy model is considered. Possible uncertainties accompanying the functioning of the system are taken into account by the inclusion in the fuzzy model of bounded external disturbances. It is shown that the problem of state estimating can be effectively solved on the basis of the technique of invariant ellipsoids.

Key words and phrases: nonlinear network control system, Takagi-Sugeno fuzzy model, state estimation, invariant ellipsoid.

УДК 519.246.8

Применение фазового анализа временных рядов для выделения макроэкономических циклов на примере динамики курсов валют

Е. А. Тевелева*, Н. П. Третьяков*[†], А. Я. Терлецкий[‡]

* Кафедра прикладных информационных технологий
Институт общественных наук
Российская академия народного хозяйства и государственной службы при
Президенте Российской Федерации
Проспект Вернадского, д. 82, стр.1, Москва, Россия, 119571

[†] Кафедра информатики и прикладной математики
Российский государственный социальный университет
ул. Вильгельма Пика, дом 4, стр.1, Москва, Россия, 129226

[‡] Институт физических исследований и технологий
Российский университет дружбы народов
ул. Миклухо-Маклая, д. 6, Москва, Россия, 117198

Email: eteveleva@yandex.ru, trn11@rambler.ru

В статье рассматривается один из методов выявления макроэкономических циклов экономических показателей на примере курса доллара за период 2000-2011гг.

Ключевые слова: макроэкономика, макроэкономический цикл, курс доллара, курс евро, анализ временных рядов.

1. Введение

Как правило, временные ряды экономических показателей, особенно за большой промежуток времени, содержат нерегулярные циклические колебания с неустойчивой амплитудой и периодом [1]. Внешне стохастический характер этих явлений отражает циклическое развитие экономики в условиях влияния на нее многих случайных и неслучайных, рыночных и волевых воздействий, то есть множества скрытых факторов, учесть которые далеко не всегда возможно. Таким образом, подобные колебания отражают макроциклы развития экономических процессов.

Исследование подобных колебаний упомянутых процессов такими методами, как например, традиционный спектральный или регрессионный анализ, вообще говоря, не подходит. Спектральный анализ моделирует движение временного ряда суммой регулярных синусоид. Однако едва ли экономические показатели обладают строгой периодичностью и постоянством амплитуд, поэтому для хорошей аппроксимации требуется большое число синусоид разной частоты. В этом случае возникают трудности с интерпретацией результатов. Регрессионный анализ аппроксимирует весь ряд в целом, не учитывая локальных свойств ряда. Между тем в экономике каждый цикл имеет свои особенности, так как порождается множеством причин самого различного характера, которые могут быть свойственны лишь определенным временным интервалам и, как правило, не повторяются. Поэтому для исследования нерегулярных циклических колебаний требуются иные методы.

2. Метод

Одним из таких методов является фазовый анализ временного ряда. Для изложения предложенного подхода введем некоторые определения [1].

Вводится понятие флуктуации (Dy_t) — величины отклонения значений ряда от некоторого установленного уровня. Это может быть отклонение от среднего значения, тренда или какой-либо другой величины. Мощность флуктуации — абсолютная величина флуктуации $|Dy_t|$. Фаза — это период положительных или отрицательных флуктуаций ряда. Длительность фазы — временной отрезок соответствующей фазы.

Таким образом, нерегулярные циклические колебания представляют собой наличие у ряда разнонаправленных флуктуаций. Мощность i -й фазы — сумма абсолютных флуктуаций ряда внутри фазы:

$$P_i = \sum_{t=t_i}^{t_i+l_i} |Dy_t|$$

где t_i — момент начала i -й фазы, t_{i+l_i} — момент окончания фазы, l_i — длительность фазы. Среднее значение фазы — мощность, усредненная на отрезке $t_i; t_{i+l_i}$.

Экономист, как правило, имеет дело с временным рядом, содержащим случайные флуктуации. В исходном ряду каждая такая флуктуация или несколько соседних могут образовывать маломощные фазы, не имеющие существенного содержания. Поэтому желательно очистить ряд от случайных флуктуаций и соответствующих им фаз, чтобы получить интерпретируемые движения ряда. Как правило, это итерационный процесс, где на каждом этапе происходит агрегирование маломощных фаз с двумя соседними, имеющими более существенную мощность. Следовательно, необходимо задать критерий остановки поглощения маломощных фаз. В качестве такого критерия могут выступать: либо уровень теряемой мощности самого ряда — заранее заданный процент совокупной мощности ряда, которую разрешается потерять в процессе агрегирования фаз (вычисляется сумма абсолютных значений — мощность — ряда, которая принимается за 100 процентов). Либо заранее заданное количество фаз. Выбор зависит от особенностей решаемой задачи. После завершения итераций строится фазовая диаграмма, где каждому моменту времени ставится в соответствие среднее значение фазы с соответствующим знаком.

Рассмотрим пример. На рис. 1 отображена динамика курса доллара за период 2000–2011 [2]. В качестве флуктуаций здесь были рассмотрены отклонения исходных значений от тренда. Тренд курса рубля (на рис. 1 он изображен пунктирной линией) статистически незначим, поэтому ось Ox соответствует примерно среднему значению доллара за указанный период, которое составило 28,75. На рис. 2 изображена фазовая диаграмма динамики курса доллара за указанный период. Каждой фазе поставлен в соответствие средний курс доллара, соответствующий каждой фазе. В качестве критерия остановки итерационного процесса здесь был принят уровень теряемой мощности элементов ряда — 5 процентов. Здесь четко выделяются три макроцикла — 2001–2003 гг., 2004–2008 гг. и цикл, начинающийся с конца 2008г.

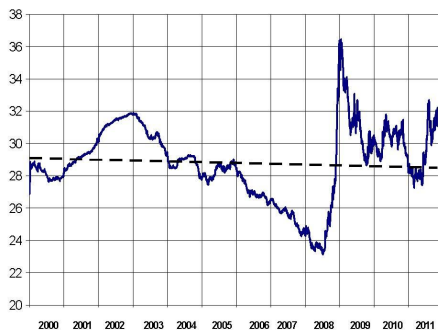


Рис. 1. Курс доллара за период 2000–2011 год

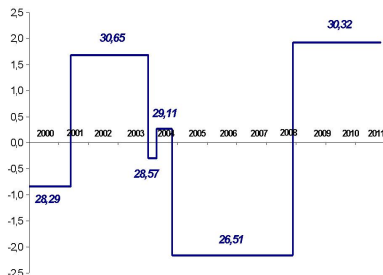


Рис. 2. Фазовая диаграмма курса доллара за период 2000–2011 год

3. Заключение

Фазовый анализ является одним из методов выявления макроэкономических циклов экономических показателей. Он имеет универсальный характер и представляет собой, по сути, фильтр, который позволяет выявить смену режимов в динамике временного ряда. Интерес представляет также сравнение результатов, полученных для нескольких рядов и обобщение метода на многомерный случай.

Литература

1. Лукашин Ю. П., Адаптивные методы краткосрочного прогнозирования временных рядов. — М.: Финансы и статистика, 2003. — 413 с.
2. URL: <http://www.kurs-dollar.ru/kurs-dollar.html>

UDC 519.246.8

Application of the Phase Analysis of Time Series for the Identification of Macroeconomic Cycles Based on the Dynamics of the Exchange Rates

E. A. Teveleva*, N. P. Tretyakov*[†], A. Ya. Terletski[‡]

* Department of Applied Information Technologies
Russian State Social University

4-1 Wilhelm Pieck str., Moscow, 119571, Russian Federation

[†] Department of Informatics and Applied Mathematics

Russian Presidential Academy of National Economy and Public Administration

82-2 Prospect Vernadskogo, Moscow, 129226, Russian Federation

[‡] Institute of Physical Research and Technologies

Peoples' Friendship University of Russia (RUDN University)

6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation

Email: eteveleva@yandex.ru, trn11@rambler.ru

The article deals with one of the methods for identifying macroeconomic cycles of economic indicators using the example of the dollar exchange rate for the period 2000–2011.

Key words and phrases: computer science, information technologies, conference proceedings.

УДК 517.958:530.145.6, 531:530.145

Эффективное уравнение Шрёдингера для одномерной системы с быстро осциллирующими граничными условиями

Н. П. Третьяков*[†], А. А. Кочанов[‡]

* Кафедра прикладных информационных технологий
Институт общественных наук

Российская академия народного хозяйства и государственной службы при
Президенте Российской Федерации

Проспект Вернадского, д. 82, стр. 1, Москва, Россия, 119571

[†] Кафедра информатики и прикладной математики
Российский государственный социальный университет
ул. Вильгельма Пика, д. 4, стр. 1, Москва, Россия, 129226

[‡] Институт физических исследований и технологий
Российский университет дружбы народов
л. Миклушко-Маклая, д. 6, Москва, Россия, 117198

Email: trn11@rambler.ru, antonkochanov@rambler.ru

В работе исследован метод эффективного уравнения Шрёдингера, примененный к одномерной задаче о движущейся частице в коробке с зависящими от времени граничными условиями.

Ключевые слова: квантовые системы с быстрым внешним воздействием, эффективное уравнение Шредингера.

1. Введение

Нахождение и исследование эффективного уравнения Шрёдингера для систем с быстрым внешним взаимодействием представляет интерес для широкого круга задач, в частности, данная задача может быть полезна при создании квантовых компьютеров нового типа. В целом нестационарные граничные условия представляют интерес ввиду возможных приложений, таких например, как атомы в полостях, ионы и атомы в магнитных ловушках, сверхпроводящие квантовые интерференционные устройства, динамический эффект Казимира. Для лучшего понимания состояния данного направления могут быть полезны следующие обзорные статьи: [2, 4, 6–9]. В данной работе исследуются системы, в которых гамильтониан (1) содержит быстро осциллирующую часть:

$$\hat{H}(\hat{x}, \hat{p}, t) = \hat{H}_0(\hat{x}, \hat{p}, t) + \sum_{n \geq 1} \hat{h}_n(\hat{x}, \hat{p}) \cos(n\Omega t) + \sum_{n \geq 1} \hat{b}_n(\hat{x}, \hat{p}) \sin(n\Omega t). \quad (1)$$

Решение (1) ищется в виде $\psi = \Phi + \mu\varphi$, где $\mu = \frac{E_0}{\hbar\Omega} \ll 1$, $E_0 = \hbar\omega_0$ — характерная энергия невозмущенной (стационарной) системы, а волновая функция $\varphi(x, t)$ «быстро» изменяется во времени (рассматривается линейное по μ приближение). Если в (1) операторы \hat{h}_n и \hat{b}_n некоммутирующие, то эффективное уравнение Шрёдингера для «медленной» волновой функции имеет вид:

$$i\hbar \frac{\partial \Phi}{\partial t} = \hat{H}_0 \Phi + \frac{1}{2i\hbar\Omega} \sum_{n \geq 1} \frac{1}{n} [\hat{b}_n(\hat{x}, \hat{p}) \hat{h}_n(\hat{x}, \hat{p}) - \hat{h}_n(\hat{x}, \hat{p}) \hat{b}_n(\hat{x}, \hat{p})] \Phi. \quad (2)$$

Уравнение (2) получено в нескольких независимых работах при помощи разных подходов (формула (33) в [1], формула (56) в [2], формула (46) в [4], формула (39) в [3], формула (16) в [5]). В настоящей работе рассматривается одномерный случай движения частицы в коробке с подвижными стенками, который, как можно увидеть ниже, сводится к уравнению (2). Необходимо отметить, что ранее метод эффективного уравнения Шрёдингера не применялся к системам с переменными граничными условиями.

2. Алгоритм вывода эффективного уравнения

Рассмотрим систему, описываемую волновой функцией ψ и представляющую собой частицу с координатой x и импульсом p , локализованную внутри одномерной коробки с координатой центра d и длиной ℓ , при этом $d = d(t)$ и $\ell = \ell(t)$. В работе [10] получено уравнение Шрёдингера для данной системы, которое можно записать следующим образом:

$$i\hbar \frac{\partial \psi}{\partial t} = (\hat{U} \hat{H}_0 \hat{U}^\dagger) \psi - \left(\frac{\dot{\ell}}{\ell} x \circ \hat{p} + \frac{\ell_0}{\ell} \dot{d} \hat{p} \right) \psi, \quad (3)$$

где H_0 — гамильтониан исходной системы, ℓ_0 — характерный размер системы, $x \circ \hat{p} = \frac{x\hat{p} + \hat{p}x}{2}$, а исходная система с зависящими от времени граничными условиями посредством унитарного оператора U преобразована в систему с постоянными граничными условиями $\psi\left(-\frac{\ell_0}{2}\right) = \psi\left(\frac{\ell_0}{2}\right) = 0$ следующим образом:

$$\psi(x) = U(\ell, d)\phi(x) = \sqrt{\frac{\ell}{\ell_0}} \phi\left(\frac{\ell}{\ell_0}x + d\right).$$

Сопряженный оператор имеет вид $U^+(\ell, d)\phi(x) = \sqrt{\frac{\ell_0}{\ell}} \phi\left(\frac{\ell_0}{\ell}x - \frac{\ell_0}{\ell}d\right)$.

Рассмотрим случай, когда частица движется в постоянном поле некоторого потенциала $V(x)$, т.е.

$$\hat{H}_0 = \frac{\hat{p}^2}{2m} + V(x).$$

Уравнение (3) после несложных преобразований принимает вид:

$$i\hbar \frac{\partial \psi}{\partial t} = \left(\frac{\ell_0}{\ell}\right)^2 \frac{\hat{p}^2}{2m} \psi + V\left(\frac{\ell}{\ell_0}x + d\right) \psi - \frac{\dot{\ell}}{\ell} x \circ \hat{p} \psi + \frac{\ell_0}{\ell} \dot{d} \hat{p} \psi.$$

Зависящие от времени граничные условия представляются в виде суперпозиции как четных, так и нечетных по времени членов

$$d(t) = d_1 \cos(\Omega t) + d_2 \sin(\Omega t); \quad \ell(t) = \ell_0 + \ell_1 \cos(\Omega t) + \ell_2 \sin(\Omega t). \quad (4)$$

Входящие в (4) амплитуды d_1, d_2, ℓ_1, ℓ_2 предполагаются малыми по сравнению с ℓ_0 .

Рассмотрим частный случай, когда ширина коробки остается постоянной, а $d(t)$ содержит только четный по времени член: $\ell_1 = \ell_2 = d_2 = 0$. Разложим потенциал по d_1/ℓ_0 до второго порядка:

$$V(x + d) \approx V(x) + \ell_0 V'(x) \frac{d_1}{\ell_0} \cos(\Omega t) + \frac{\ell_0^2}{2} V''(x) \left(\frac{d_1}{\ell_0}\right)^2 \cos^2(\Omega t).$$

Уравнение (3) приводится тогда к виду (1):

$$i\hbar \frac{\partial \psi}{\partial t} = \left(\frac{\hat{p}^2}{2m} + V(x) + \frac{d_1^2}{4} V''(x)\right) \psi + (d_1 V'(x) - \Omega d_2 \hat{p}) \cos(\Omega t) + (d_2 V'(x) + \Omega d_1 \hat{p}) \sin(\Omega t) + \frac{d_1^2}{4} V''(x) \cos(2\Omega t). \quad (5)$$

Подстановка коэффициентов h_1, b_1, h_2 , возникших в (5), в (2) приводит к эффективному уравнению Шрёдингера:

$$i\hbar \frac{\partial \Phi}{\partial t} = \left(\frac{\hat{p}^2}{2m} + V(x)\right) \Phi - \frac{d_1^2}{4} V''(x) \Phi. \quad (6)$$

3. Заключение

В рассмотренном частном примере, как видно из (6), эффект быстрого колебания центра коробки свелся в эффективном уравнении к перенормировке потенциальной энергии. Разработанный алгоритм может быть использован для вывода эффективных уравнений в других случаях (колебания стенок коробки, колебания параметров потенциала), в которых могут возникать более интересные эффекты.

Литература

1. Novičenko V., Anisimovas E., and Juzeliūnas G., Floquet analysis of a quantum system with modulated periodic driving // *Phys. Rev. A.* — 2017. — Vol. 95. — P. 023615.
2. Weinberg Ph., Bukov M., D'Alessio L., Polkovnikov A., Vajna S., Kolodrubetz M. Adiabatic Perturbation Theory and Geometry of Periodically-Driven Systems // *Physics Reports.* — 2017. — Vol. 688, 1
3. Tretyakov N. P. and Aguero M. A. Effective Schrödinger equation for fast driven quantum systems // *Phys. Scr.* — 2015. — Vol. 90. — P. 085207.
4. Bukov M., D'Alessio L. and Polkovnikov A. Universal High-Frequency Behavior of Periodically Driven Systems: from Dynamical Stabilization to Floquet Engineering // *Advances in Phys.* — 2015. — Vol. 64. — P. 139.
5. Goldman N., Dalibard J. Periodically Driven Quantum Systems: Effective Hamiltonians and Engineered Gauge Fields // *Physical Review.* — 2014. — X 4. — P. 031027
6. Silveri M. P., Tuorila J. F., Thuneberg E. V., Paraoanu G. S. Quantum systems under frequency modulation // *Rep. Prog. Phys.* — 2017. — Vol. 80(5). — P. 056002
7. Gritsev V., Polkovnikov A. Integrable Floquet Dynamics // *SciPost Phys.* — 2017. — Vol. 2 (021).
8. Hanggi P. Driven Quantum Systems, 2016.
9. Rahav S., Gilary I., Fishman S. Effective Hamiltonians for periodically driven systems // *Phys. Rev. A.* — 2003. — Vol. 68. — P. 013820.
10. Di Martino S., Facchi P. Quantum systems with time-dependent boundaries, Preprint arXiv: 1501.06490v1 [math-ph] (2015).

UDC 517.958:530.145.6, 531:530.145

Effective Schrödinger Equation for One-Dimensional Systems with Rapidly Oscillating Boundary Conditions

N. P. Tretyakov*[†], A. A. Kochanov[‡]

* *Department of Applied Information Technologies
Russian State Social University*

4 Wilhelm Pieck str., build.1, Moscow, 119571, Russian Federation

[†] *Department of Informatics and Applied Mathematics*

*Russian Presidential Academy of National Economy and Public Administration
Prospect Vernadskogo, 82, bldg 2, Moscow, 129226, Russian Federation*

[‡] *Institute of Physical Research and Technology,*

*Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: trn11@rambler.ru, antonkochanov@rambler.ru

The method of the effective Schrödinger equation applied to one-dimensional problems of particles in a potential box with high-frequency time-dependent boundary conditions is investigated.

Key words and phrases: quantum systems with fast external action, the effective Schrödinger equation.

УДК 519.6

Устойчивое вычисление нормали к поверхности, заданной приближенно**К. В. Шебанова, К. И. Хамидуллина***Российский университет дружбы народов
ул. Миклухо-Маклая, д. 6, Москва, Россия, 117198*

Email: akayntas@mail.ru, karina-oren@mail.ru

В работе предложен устойчивый метод построения нормали к неточно заданной поверхности.

Ключевые слова: нормаль и поверхности, некорректные задачи.

1. Введение

При решении многих прикладных задач математической физики требуется вычисление нормали к поверхности [1]. Вычисление нормали связано с вычислением производных от функции в уравнении поверхности, а вычисление производных — некорректно поставленная задача [2], и если поверхность задана неточно, то это может приводить к неустойчивости в вычислении нормали.

2. Постановка задачи.

Пусть имеется незамкнутая поверхность S в цилиндре D^∞ прямоугольного сечения

$$D^\infty = \{(x, y, z) : 0 < x < l_x, 0 < y < l_y, -\infty < z < \infty\}.$$

Будем считать, что эта поверхность задается уравнением

$$F(x, y, z) = 0, \quad (x, y, z) \in D^\infty.$$

Как известно [3], нормаль (не нормированная) вычисляется как градиент функции F :

$$\vec{n} = \text{grad}F(x, y, z) = \left\{ \frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial F}{\partial z} \right\}.$$

Пусть теперь вместо точной функции F известна функция F^δ , такая, что

$$\|F^\delta - F\|_{L_2(D^\infty)} = \delta.$$

Приближенный вектор нормали будем находить как градиент функционала Тихонова

$$M^\alpha[w] = \|w - F^\delta\|_{L_2(D^\infty)}^2 + \alpha \|grad w\|_{L_2(D^\infty)}^2.$$

Экстремаль w_α^δ функционала $M^\alpha[w]$ является решением уравнения Эйлера для этого функционала

$$-\alpha \Delta w + w = F^\delta.$$

и имеет непрерывные производные. Экстремаль получена методом Фурье. Для приближенного вектора нормали получаем

$$\vec{n}_\alpha^\delta = grad w_\alpha^\delta.$$

3. Выводы

При согласовании параметра регуляризации α с погрешностью δ в задании функции F сходимость приближенного вектора нормали к точной нормали на поверхности S :

$$\|\vec{n}_\alpha^\delta - \vec{n}\|_{L_2(S)} \rightarrow 0 \quad \text{при} \quad \delta \rightarrow 0.$$

Литература

1. Ланеев Е. Б. Некорректные задачи продолжения гармонических функций и потенциальных полей и методы их решения. — М. Изд-во РУДН, 2006.
2. Тихонов А. Н., Арсенин В. Я. Методы решения некорректных задач. — М.: Наука, 1979.
3. Ильин В. А., Позняк Э. Г. Основы математического анализа, часть 1. — М.: Наука, 1971.

UDC 519.6

Stable Calculations of Normal to the Surface with Set Approximately

K. V. Shebanova, K. I. Khamidulina

*Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: akayntas@mail.ru, karina-oren@mail.ru

A stable method for constructing a normal to an inaccurately defined surface is proposed in this paper.

Key words and phrases: normal and surfaces, incorrect tasks.

УДК 519.872

Моделирование прохождения световой волны через открытый планарный волновод методом «ящика Дирихле»

В. К. Чупритский, Д. В. Диваков

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: vkchupritskij@sci.pfu.edu.ru, divakov_dv@rudn.university

В работе описывается метод расчёта направляемых и излучательных волноводных мод в открытых тонкоплёночных планарных волноводах, помещенных в «ящик Дирихле».

Ключевые слова: математическое моделирование, открытые волноводы, излучательные моды, направляемые моды.

1. Введение

В работе рассматривается планарный волновод, в котором имеется три слоя, средний из которых, называемый волноводным, представляет собой тонкую плёнку толщиной в несколько длин волн, он характеризуется показателем преломления n_f и заключен в области пространства $x \in (0, h)$. Он нанесён на подложку, которая описывается показателем преломления n_s . Над волноводным слоем находится покровный слой, характеризующийся показателем преломления n_c . Для эффективного прохождения света вдоль волноводного слоя подбираются материалы таким образом, чтобы выполнялось следующее соотношение для показателей преломления: $n_c < n_s < n_f$.

2. Основная часть

Свет представляет собой электромагнитное поле, его распространение в волноводе описывается уравнениями Максвелла [1]. В планарных структурах уравнения Максвелла разбиваются на уравнения Гельмгольца для TE - и TM -мод и две пары дополнительных соотношений. В данной работе рассматривается TE мода, которая описывается компонентой E_y электрического поля и двумя компонентами магнитного, при этом компонента E_y удовлетворяет уравнению Гельмгольца, приведенному ниже:

$$\left(\frac{\partial^2}{\partial z^2} + \frac{\partial^2}{\partial x^2} + k_0^2 n^2(x) \right) E_y = 0. \quad (1)$$

где $n^2(x)$ – показатель преломления, k_0 – волновое число в вакууме.

В открытых волноводах могут распространяться направляемые моды [1]. Энергия распространяющегося излучения в этом случае сосредоточена в волноводном слое и области, непосредственно примыкающей к нему. Кроме направляемых мод в открытых волноводах могут также возникать так называемые излучательные моды. Для излучательных мод энергия не локализована в непосредственной близости от волноводного слоя, а распределена по всей оси. Адекватный учёт излучательных мод в задачах дифракции основан на дискретизации непрерывного спектра. В рамках настоящей работы использована идея так называемого «ящика Дирихле» [3], состоящая в помещении рассматриваемого открытого волновода в объемлющий полый закрытый волновод с границами $x = \pm R_x$, на которых выполняются условия Дирихле:

$$\begin{cases} E_y |_{x=-R_x} = 0, \\ E_y |_{x=+R_x} = 0. \end{cases} \quad (2)$$

На границах раздела сред должны выполняться условия непрерывности тангенциальных компонент электромагнитного поля, которые имеют следующий вид для TE -мод [2]:

$$[E_y]_{x=0,h} = \left[\frac{\partial E_y}{\partial x} \right]_{x=0,h} = 0, \quad (3)$$

где $[f]_{x=a}$ — скачок функции f в точке $x = a$.

Воспользуемся методом разделения переменных, а именно представим искомую величину E_y в виде произведения функций от переменных x и z :

$$E_y(x, z) = \psi(x) \cdot V(z).$$

Применив метод разделения переменных и произведя преобразования уравнения (1) и граничных условий (2) и (3), получим задачу Штурма–Лиувилля [3]:

$$\begin{cases} \psi'' + k_0^2 (n^2(x) - \beta^2) \psi = 0, \\ \psi(\pm R_x) = 0, \quad [\psi]_{x=0,h} = 0, \quad [\psi']_{x=0,h} = 0. \end{cases} \quad (4)$$

Задача будет решаться для следующих условий: $\lambda = 0.55\text{мкм}$, $h = 2\lambda$, $R_x = 8\lambda$, $n_f = 1.55$, $n_s = 1.47$, $n_c = 1$.

Для решения задачи (4) воспользуемся методом волнового сопряжения: в областях постоянства $n^2(x)$ запишем общее решение уравнения из (4), в которое входят неопределённые константы и параметр β . Подставив вид решения в граничные условия из (4), получим однородную систему линейных алгебраических уравнений вида $M(\beta)\vec{a} = \vec{0}$. Из условия существования нетривиального решения $\det M(\beta) = 0$ отыщем собственные значения β_j при помощи метода дихотомии. Для каждого β_j далее решим систему $M(\beta_j)\vec{a}_j = \vec{0}$ и, используя модифицированный метод Гаусса с выбором главного элемента, отыщем неизвестные константы, которые определяют соответствующие собственные функции $\psi_j(x)$ (графики некоторых из них приведены на рис. 1–3).

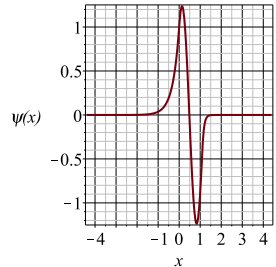
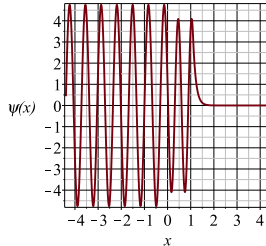
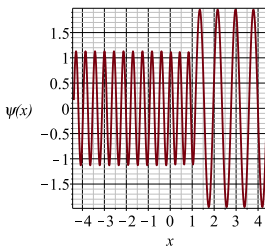


Рис. 1. График функции $\psi(x)$ при $\beta = 1.49$

Рис. 2. График функции $\psi(x)$ при $\beta = 1.23$

Рис. 3. График функции $\psi(x)$ при $\beta = 0.74$

3. Заключение

В работе описан метод расчёта волноводных мод, распространяющихся в открытом тонкоплёночном планарном волноводе, помещенном в «ящик Дирихле», построены графики поперечного распределения интенсивности волноводных мод. Данный метод в дальнейшем будет использоваться для численного решения задачи дифракции волновых мод на стыке двух планарных волноводов и последующих прикладных задач.

Литература

1. Самарский А. А., Тихонов А. Н. О представлении поля в волноводе в виде суммы полей ТЕ и ТМ // ЖТФ. — 1948. — Т. 18, вып. 7. — С. 959–970.
2. Могилевский И. Е., Свешников А. Г. Математические проблемы теории дифракции. — М.: МГУ, 2010.
3. Диваков Д. В., Малых М. Д., Севастьянов А. Л., Севастьянов Л. А. Моделирование распространения поляризованного света в тонкоплёночной волноводной линзе // Вестник РУДН. Серия: Математика. Информатика. Физика. — 2017. — Т. 25, № 1. — С. 56–68.

UDC 519.872

Simulation of Light Propagation Through Open Planar Waveguide with the Use of “Dirichlet Box” Method

V. K. Chupritskiy, D. V. Divakov

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russia*

Email: vkchupritskij@sci.pfu.edu.ru, divakov_dv@rudn.university

This paper describes the process of calculation of transversal distribution of waveguiding modes intensity which come through open thin-film planar waveguide placed in “Dirichlet box”.

Key words and phrases: mathematical modeling, open waveguides, radiative modes, guided modes.

UDC 519.86

On Hyperbolic-Sine Local Volatility Model

S. G. Shorokhov, A. E. Buurul dai

*Nikol'skii Mathematical Institute
Peoples' Friendship University of Russia
Miklukho-Maklaya str. 6, Moscow, 117198, Russian Federation*

Email: shorokhov_sg@rudn.university, aivo4ka@mail.ru

We study a local volatility model based on stochastic process of hyperbolic-sine type. We derive the transition probability density function for hyperbolic-sine model and justify that this function has delta function terminal condition at initial time. Risk neutral valuation technique is applied to find explicit valuation formula for european call option in hyperbolic-sine model.

Key words and phrases: stochastic models, volatility function, option pricing.

1. Introduction

Modern mathematical finance is based on the famous constant volatility Black-Scholes model [1]. But an assumption of constant volatility fails to hold in practice because of the so-called “volatility smiles” and fat tails of financial data distributions. One of the most natural approaches to these issues is the transition to general model of risk-neutral asset dynamics of the form

$$dS = r S dt + \sigma(S, t) S, \quad dW, \quad S(t_0) = S_0, \quad (1)$$

with volatility σ being a function of asset price S and time t .

Known examples of local volatility models include shifted lognormal model [2] with volatility function $\sigma_0 \left(1 - \frac{\alpha e^{r t}}{S}\right)$, $\alpha \in \mathbb{R}$, normal (Cox-Ross) model [3] with volatility function $\frac{\sigma_0}{S}$ and CEV model [3] with volatility function $\sigma_0 S^\beta$.

2. Hyperbolic-Sine Local Volatility Model

Let the dynamics of the asset price S under some risk-neutral measure be given by stochastic differential equation (SDE)

$$dS = r S dt + \sigma_0 \sqrt{1 + \frac{2r}{\sigma_0^2} S^2} dW, \quad (2)$$

where $r > 0$, $\sigma_0 > 0$, W — standard Brownian motion, $S(t_0) = S_0 > 0$. Model (2) implies underlying stochastic process of hyperbolic-sine type, which was proposed in [4].

Changing variable S in (2) to $S' = \text{arsinh}\left(\frac{\sqrt{2r}}{\sigma_0} S\right) = \ln\left(\frac{\sqrt{2r}}{\sigma_0} S + \sqrt{1 + \frac{2r}{\sigma_0^2} S^2}\right)$ and applying Ito's Lemma we obtain the following SDE for new variable S'

$$dS' = \sqrt{2r} dW \quad (3)$$

with initial condition $S'(t_0) = S'_0 = \text{arsinh}\left(\frac{\sqrt{2r}}{\sigma_0} S_0\right)$. From (3) it follows that the stochastic process $S'(t)$ is distributed normally with expectation $S'_0 = \text{arsinh}\left(\frac{\sqrt{2r}}{\sigma_0} S_0\right)$ and variance $2r(t - t_0)$, i.e.

$$S'(t) \sim \mathcal{N} \left(\operatorname{arsinh} \left(\frac{\sqrt{2r}}{\sigma_0} S_0 \right), 2r(t-t_0) \right).$$

The transition probability density function for $S'(t)$ is equal to

$$\rho'(x', t, S'_0, t_0) = \frac{1}{2\sqrt{\pi r(t-t_0)}} e^{-\frac{(x'-S'_0)^2}{4r(t-t_0)}}, \quad t \geq t_0.$$

The relation $S(t) = \frac{\sigma_0}{\sqrt{2r}} \sinh S'(t)$ implies that the transition probability density function for $S(t)$ is equal to

$$\begin{aligned} \rho(x, t, S_0, t_0) &= \rho' \left(\operatorname{arsinh} \left(\frac{\sqrt{2r}}{\sigma_0} x \right), t, \operatorname{arsinh} \left(\frac{\sqrt{2r}}{\sigma_0} S_0 \right), t_0 \right) \frac{d}{dx} \operatorname{arsinh} \left(\frac{\sqrt{2r}}{\sigma_0} x \right) = \\ &= \frac{1}{\sqrt{2\pi} \sqrt{\sigma_0^2 + 2rx^2} \sqrt{t-t_0}} e^{-\frac{1}{4r(t-t_0)} \left(\operatorname{arsinh} \left(\frac{\sqrt{2r}}{\sigma_0} x \right) - \operatorname{arsinh} \left(\frac{\sqrt{2r}}{\sigma_0} S_0 \right) \right)^2}. \quad (4) \end{aligned}$$

Function ρ , given by (4), is a solution of Fokker–Planck partial differential equation (PDE) for hyperbolic-sine model (2):

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x} (rx\rho) - \frac{1}{2} \frac{\partial^2}{\partial x^2} ((\sigma_0^2 + 2rx^2)\rho) = 0.$$

The expectation of $S(t)$ is equal to

$$\mathbb{E}[S(t)] = \int_{-\infty}^{+\infty} x\rho(x, t, S_0, t_0) dx = S_0 e^{r(t-t_0)}.$$

The variance of $S(t)$ is equal to

$$\begin{aligned} \mathbb{V}[S(t)] &= \mathbb{E}[S^2(t)] - (\mathbb{E}[S(t)])^2 = \int_{-\infty}^{+\infty} x^2 \rho(x, t, S_0, t_0) dx - S_0^2 e^{2r(t-t_0)} = \\ &= S_0^2 e^{4r(t-t_0)} + \frac{\sigma_0^2}{4r} (e^{4r(t-t_0)} - 1) - S_0^2 e^{2r(t-t_0)}. \end{aligned}$$

Obviously, if $t \rightarrow t_0$ then $\mathbb{E}[S(t)] \rightarrow S_0$, $\mathbb{V}[S(t)] \rightarrow 0$, which means that $S(t_0)$ has Dirac delta distribution concentrated at the point S_0 .

In hyperbolic-sine model (2) european call option price $c(S, t, K, T)$ satisfies Black–Scholes–Merton PDE

$$\frac{\partial c}{\partial t} + rS \frac{\partial c}{\partial S} + \frac{1}{2} (\sigma_0^2 + 2rS^2) \frac{\partial^2 c}{\partial S^2} - rc = 0$$

with boundary condition $c(S, T, K, T) = \max(S - K, 0)$ and can be determined as

$$c(S, t, K, T) = e^{-r(T-t)} \mathbb{E}[\max\{S - K, 0\}] = e^{-r(T-t)} \int_K^{+\infty} (x - K) \rho(x, T, S, t) dx. \quad (5)$$

Changing variable x in (5) to $u = \frac{1}{\sqrt{2r(T-t)}} \left(\operatorname{arsinh} \left(\frac{\sqrt{2r}}{\sigma_0} x \right) - \operatorname{arsinh} \left(\frac{\sqrt{2r}}{\sigma_0} S \right) \right)$ we calculate european call option price as follows

$$c(S, t, K, T) = \frac{1}{2} S \left(\Phi \left(-K^* - \sqrt{2r(T-t)} \right) + \Phi \left(-K^* + \sqrt{2r(T-t)} \right) \right) + \\ + \frac{1}{2} \sqrt{\frac{\sigma_0^2}{2r}} + S^2 \left(\Phi \left(K^* + \sqrt{2r(T-t)} \right) - \Phi \left(K^* - \sqrt{2r(T-t)} \right) \right) e^{-r(T-t)} K \Phi(-K^*),$$

where $K^* = \frac{1}{\sqrt{2r(T-t)}} \left(\operatorname{arsinh} \left(\frac{\sqrt{2r}}{\sigma} K \right) - \operatorname{arsinh} \left(\frac{\sqrt{2r}}{\sigma} S \right) \right)$.

3. Conclusion

We presented a hyperbolic-sine local volatility model including the transition probability density function as an exact solution of Fokker–Planck PDE and the european call option pricing formula as an exact solution of Black–Scholes–Merton PDE. Hyperbolic-sine local volatility model captures some aspects of the volatility smiles and can be used for pricing different types of options. The model can also be used for credit risk estimation in structural credit risk models.

References

1. F. Black, M. Scholes, The Pricing of Options and Corporate Liabilities, The Journal of Political Economy 81 (3), pp. 637–654.
2. D. Brigo, F. Mercurio, Fitting volatility skews and smiles with analytical stock-price models, Seminar Paper at Institute of Finance, University of Lugano, 2000. Available on the Internet at <http://www.istfin.eco.unisi.ch/seminar-papers-smile.pdf>.
3. J. C. Cox, S. A. Ross, The valuation of options for alternative stochastic processes, Journal of Financial Economics, 3 (1–2), 145–166. [https://doi.org/10.1016/0304-405X\(76\)90023-4](https://doi.org/10.1016/0304-405X(76)90023-4)
4. D. Brigo, F. Mercurio, G. Sartorelli, Alternative asset-price dynamics and volatility smile, Quantitative Finance 3 (3), 173–183.

УДК 519.86

О модели локальной волатильности с гиперболическим синусом

С. Г. Шорохов, А. Э. Буурулдай

*Математический институт им. С.М. Никольского
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: shorokhov_sg@rudn.university, aivo4ka@mail.ru

Рассматривается модель локальной волатильности, основывающаяся на стохастическом процессе с гиперболическим синусом. Для рассматриваемой модели выводится переходная плотность вероятности и проверяется, что начальное условие для переходной плотности вероятности представляет собой дельта-функцию Дирака. При помощи риск-нейтрального подхода к ценообразованию получена явная формула стоимости европейского опциона колл в модели с гиперболическим синусом.

Ключевые слова: стохастические модели, функция волатильности, ценообразование опционов.

УДК 519.68:159.955

Сравнение Online-Q и Experience Replay для обучения нейронной сети в мультиагентной среде

Е. В. Шунин

*Математический институт имени С.М. Никольского
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: shunin_e@mail.ru

В данной работе проводится сравнение двух методов обучения с подкреплением, а именно Online-Q и Experience replay. Сравнение проведено для агента, обучаемого в мультиагентной среде, которая представляет собой видеоигру “mitos.is”. При использовании нейронной сети в обучении с подкреплением возникают трудности с высокой корреляцией между примерами, которая осложняет обучение агента. Таким образом использование experience replay позволяет работать с независимыми примерами. В ходе проведенного сравнения выявлено преимущество experience replay для данной среды.

Ключевые слова: нейронные сети, обучение с подкреплением, Q обучение, информационные технологии, проведение конференции.

1. Введение

Обучение с подкреплением — это подход к тому, чтобы произвести отображение из пространства состояний в пространство действий с учетом максимизации получаемого вознаграждения [1]. Предположим, что у нас имеется некоторая среда и агент взаимодействующий с ней. За каждое действие, предпринятое агентом, среда поощряет его наградой. Цель агента найти оптимальную стратегию поведения в данной среде, чтобы максимизировать получаемую награду.

Введем следующий вектор (s, a, r, s') , где s — состояние в момент времени t , a — действие выполненное агентом в состоянии s , r — награда полученная в результате выполнения действия a , s' — состояние, в которое переходит агент после выполнения действия a .

Для выбора агентом наиболее выгодного действия, введем понятие ценности состояние-действие $Q(s, a)$. Уточнение функции ценности будем производить по следующему правилу [1]:

$$Q(s_t, a_t) = (1 - \alpha) * Q(s_t, a_t) + \alpha * (r_t + \gamma * \max Q(s_t + 1, a)).$$

В данной работе проводится сравнение результатов полученных при обучении агента методами online-Q и experience replay. Под обучением агента будем понимать процесс аппроксимации функции ценности нейронной сетью [3]. В качестве среды воспользуемся видеоигрой mitos.is. Mitos.is представляет собой мультиагентную среду. Это означает, что агент взаимодействует не только со средой, но и с другими игроками.

Аппроксимация функции ценности нейронной сетью в задачах обучения с подкреплением позволяет построить агента в случае, когда нет информации о всех возможных состояниях среды или количество таких состояний велико.

Применение нейронной для задач обучения с подкреплением имеет две основные проблемы: Во-первых, требуется большое количество данных для обучения. Во-вторых, примеры на которых обучается нейронная сеть не являются независимыми и коррелируют между собой, что приводит к «забыванию» сетью части примеров и переобучению на последних примерах [2].

2. Сравнение online-Q и Experience replay

Online-Q: уточнение оценки полученной от нейронной сети происходит в ходе взаимодействия агента со средой. После выполнения действия a , агент получает награду r и переходит в следующее состояние s' . Далее оценка сети уточняется на векторе (s, a, r, s') .

Experience replay: в этом случае, оценка уточняется не на одном примере, а на случайной выборке векторов вида (s, a, r, s') , накопленных в ходе взаимодействия агента со средой [3].

Для сравнения этих методов построим агента для взаимодействия со средой, которая реализована в видеоигре *mitos.is*. Задача агента выбрать направление движения, таким образом чтобы максимизировать свою награду в ходе игры. Здесь s — изображение на экране, a — направление, в котором будет двигаться агент. В случае, если агент проигрывает на него накладывается штраф.

Для аппроксимации функции ценности построим нейронную сеть, которая на вход принимает 6 изображений, соответствующих состоянию игры за последнюю секунду. Выход сети представляет собой вектор y размерности k , где k — количество возможных действий агента. Агент реализует жадную стратегию это означает, что выбирается действие с наибольшей ценностью, то есть $\operatorname{argmax}(y)$ [1, 2].

Архитектура сети представляет собой сеть прямого распространения с 3 сверточными и 3 полносвязными слоями. Параметры сети инициализированы случайным образом из отрезка $(0, 1)$. Это также приводит к тому, что первое приближение Q выбирается случайно.

В дальнейшем было произведено обучение сети с помощью online-Q и experience replay. Во втором случае уточнение оценки происходило каждые 10 игр.

Как видно из графика при обучении сети с experience replay, поведение агента быстрее сходится к оптимальному, чем при использовании online-Q.

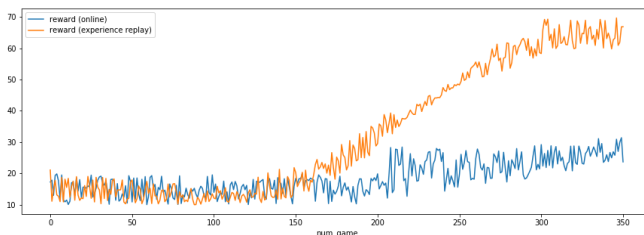


Рис. 1. Кривая обучения агента. По оси y — средняя награда достигнутая агентом в ходе игры, по оси x — номер группы из 10 игр, по которым рассчитывалась средняя награда

3. Заключение

Исходя из проведенного опыта и кривой обучения видно, что агент тратит больше времени на поиск оптимального поведения в заданной среде при использовании online-Q. Это обусловлено переобучением нейронной сети на более поздних примерах.

Аналогичный эксперимент проводился для игр Atari [2], которые представляют только взаимодействие среды и агента. В данном случае применение experience replay для мультиагентной среды также показывает преимущество над использованием online-Q.

Литература

1. Sutton R. S., Barto A. G. Reinforcement Learning: An Introduction, MIT Press, Cambridge, MA, 1998.
2. Mnih V., Kavukcuoglu K., Silver D., Graves A., Antonoglou I., Wierstra D., Riedmiller M. Playing Atari with Deep Reinforcement Learning <https://arxiv.org/abs/1312.5602>.
3. Zhang S., Sutton R. S. A Deeper Look at Experience Replay <https://arxiv.org/abs/1712.01275>.
4. Goodfellow I., Bengio Y., Courville A. Deep Learning. — MIT Press, Cambridge, MA, 2016.

UDC 519.68:159.955

Comparison of “Online-Q” and “Experience Replay” for Training a Neural Network in a Multi-Agent Environment

E. V. Shunin

*Nicol'skii Mathematical Institute
Peoples' Friendship University of Russia
Miklukho-Maklaya str. 6, Moscow, 117198, Russia*

Email: shunin_e@mail.ru

In this paper, two training methods with reinforcement learning are compared, namely Online-Q and Experience replay. A comparison was made for an agent trained in a multi-agent environment, which is a video game "mitos.is". When using a neural network in reinforced learning, difficulties arise with a high correlation between the examples, which complicates the training of the agent. Thus, using experience replay gives independent examples. Comparison reveals the advantage of experience replay for this environment.

Key words and phrases: neural networks, reinforcement learning, Q-learning, information technologies, conference proceedings.

UDC 519.234

On Some Properties of Tail Dependence Coefficient Nonparametric Estimators

Eu. Yu. Shchetinin

Financial Unviersity

*Department of Data Analysis, Decision Making and Financial Technologies
Leningradsky pr. 49, Moscow, 117324, Russian Federation*

Email: riviera-molto@mail.ru

Extreme rainfall events are of particular importance due to their severe impacts on the economy, the environment and the society. Characterization and quantification of extremes and their spatial dependence structure may lead to a better understanding of extreme events. An important concept in statistical modeling is the tail dependence coefficient (TDC) that describes the degree of association between concurrent rainfall extremes at different locations. The knowledge of the spatial characteristics of the TDC can help improve on the existing models of the occurrence probability of extremes. In this study, efficient estimation of the TDC in rainfall is investigated. Different nonparametric TDC estimators are implemented on samples of the rain gauge data and their advantages and disadvantages are discussed. The results indicate that a significant tail dependency may exist that cannot be ignored for realistic modeling of multivariate rainfall fields. Presence of a strong dependence among extremes contradicts with the assumption of joint normality, commonly used in hydrologic applications.

Key words and phrases: extreme rainfall, tail dependence coefficient, nonparametric estimator, copula.

1. Introduction

Characterization and quantification of extremes and their spatial dependence structure may lead to better estimates of probability occurrence of rare events. Most commonly used measures of dependence such as the Pearson linear correlation and Spearman correlation are not able to correctly describe the dependence of extremes. While the Spearman correlation always exists, the Pearson linear correlation may not exist for random variables above a certain extreme threshold. In general, most measures of dependence are based on the association of the entire distributions of multiple variables. However, the degree of association (dependence) between extreme values may be significantly different than that of the mid-range values. An important concept in extreme value analysis is the tail dependence coefficient (TDC) which describes the dependence in the tail of a multivariate distribution [1, 2]. S. Coles [3] introduced the TDC as the degree of association in the upper-right quadrant and lower left quadrant of a bivariate distribution. In a bivariate distribution function the tail dependence describes the limiting proportion that one marginal distribution exceeds a given threshold conditioned on the fact that the other margin has already exceeded that threshold.

2. Nonparametric Estimators of the Tail-Dependence

In this paper we consider nonparametric estimation of TDC as it's much more general approach and no assumptions are made about the marginal distribution functions and dependence structure. According to Sklar's fundamental theorem, there exists a unique decomposition

$$F_{X,Y}(F_X(x), F_Y(y)) = C(F_X(x), F_Y(y))$$

of the joint distribution $F_{X,Y}(x, y)$ and its marginal distribution functions $F_X(x)$, $F_Y(y)$ into the so-called copula function [2]:

$$C(u, v) = P(U \leq u, V \leq v), \quad (1)$$

where $U = F_X(X)$, $V = F_Y(Y)$ are uniform random variables on $[0, 1]^2$. The upper tail dependence coefficient (TDC) is usually defined as

$$\lambda_U = \lim_{u \rightarrow 1^-} P \left\{ Y > F_Y^{-1}(u) \mid X > F_X^{-1}(u) \right\} = \lim_{u \rightarrow 1^-} \frac{1 - 2u + C(u, u)}{1 - u}. \quad (2)$$

S. Coles [3] provides asymptotically equivalent version of (2)

$$\lambda_U = 2 - \lim_{u \rightarrow 1^-} \frac{\log C(u, u)}{u}.$$

For order statistics from pairs (X_n, Y_n) let $X_{(1)} \equiv \min(X_1, \dots, X_n) \leq \dots \leq X_{(n)} \equiv \max(X_1, \dots, X_n)$. Then non-parametric copula estimator took place

$$C_n \left(\frac{i}{n}, \frac{j}{n} \right) = \sum_{l=1}^n I \{ X_l \leq X_{(i)}, Y_l \leq Y_{(j)} \}. \quad (3)$$

Plugging (2) into (3) we can get well known estimator $\lambda_U^{(LOG)}$ [3]:

$$\lambda_U^{(LOG)} = 2 - \frac{\log(C_n(1 - k/n, 1 - k/n))}{\log(1 - k/n)},$$

where $k \approx \sqrt{n}$ seems to be appropriate. Next, we propose a new TDC nonparametric estimator λ_U^{RG} as a solution of regression equation

$$\log C_n \left(1 - \frac{i}{n}, 1 - \frac{i}{n} \right) = (2 - \lambda_U) \log \left(\frac{i}{n} \right) + \varepsilon_i, \quad (4)$$

where $C^*(u, v) = (\min(u, v))^\delta (uv)^\delta$, $\delta \in [0, 1]$ is a suitable model of copula (1). To compare the tail dependence models a plateau-finding algorithm is used to obtain TDC independent of a fix threshold [5]. Using this method, for each pair of rainfall data the TDC are estimated based on the optimal threshold. The results of plateau-finding algorithm showed that estimation of TDC independent of a fixed threshold is superior to the threshold-based approach.

3. Conclusions

The statistical analysis of rainfall extremes is of particular importance in risk assessment and decision making. Additionally, extreme events and their spatial dependencies are important for practical hydrologic applications such as characterization of intense rainfall events and simultaneous floods. The concept of tail dependence is commonly used to describe the degree of association in the upper tail of a multivariate distribution. This study surveys four nonparametric tail dependence approaches implemented on rainfall time series of the european part of Russia in 1966–2017 years.

The nonparametric methods are defined based on the bivariate empirical copula of pairs of variables have been compared using simulations with five most popular families of copulas (Clayton, Frank, Gumbel, normal, mixed copula (4)). The aim of this comparison was to choose the best estimator in the context of our application.

The two best estimators in our case are λ_U^{LOG} and λ_U^{RG} . The latter has the advantage that no threshold is needed.

The analysis of TDC over different temporal durations, shows that the longer duration data the higher TDC, which is consistent with the findings of [4, 5]. The performed

inter-gauge rainfall analyses show that significant tail dependency may exist that cannot be ignored. However, numerous copula models (Gaussian and meta-Gaussian models) ignore the presence of tail dependence. Further studies over different temporal and spatial resolutions are required to characterize the tail dependence coefficient for practical applications.

References

1. P. Embrechts, Kluppelberg C. and Mikosch T., Modelling Extremal Events, Springer Verlag, Berlin, 1997
2. H. Joe, Multivariate Models and Dependence Concepts, Chapman and Hall, London, 1999.
3. S. Coles, J. Heffernan, J. Tawn, Dependence measures for extreme value analyses. Extremes, 1999, 2, 339–365.
4. J. Dobric, F. Schmid, Nonparametric Estimation of the Lower Tail Dependence λ in Bivariate Copulas. Journal of Applied Statistics, 2005, 32(4), 387–407.
5. G. Frahm, M. Junker, R. Schmidt, Estimating the tail dependence coefficient. Insurance: Mathematics and Economics, 2005, 37, 80–100.

УДК 519.234

О некоторых свойствах непараметрических оценок коэффициентов хвостовой зависимости

Е. Ю. Щетинин

*Финансовый Университет при Правительстве Российской Федерации
Департамент анализа данных, принятия решений и финансовых технологий
Ленинградский пр., д. 49, Москва, Россия, 117324*

Email: riviera-molto@mail.ru

Экстремальные осадки представляют собой серьезную угрозу экономике, окружающей среде и обществу. Важной характеристикой структур статистической зависимости является коэффициент хвостовой зависимости (TDC), описывающий силу связи между экстремумами осадков. Численные эксперименты по моделированию полей осадков на территории центральной части России за 1966–2017 г.г. показали наличие значительной хвостовой зависимости между экстремумами в различных их локациях. В настоящей работе проведены исследования свойств непараметрических оценок TDC, рассмотрены их преимущества и недостатки, а также предложена новая непараметрическая оценка, использующая представление модели структур статистических связей экстремумов в виде композиции копул, что позволило получить ее в виде решения уравнения линейной регрессии. В работе проведен анализ оценок TDC для различных длительностей осадков (от 1 часа до 24 часов), показавший, что с ростом длительности величина TDC возрастает, что согласуется с выводами [4]. Также показано, что существует зависимость TDC от расстояния между локациями источников осадков. Для ее численного моделирования предложено использовать смешанную регрессионную модель, параметры которой подбираются по всем длительностям осадков методом условного правдоподобия.

Ключевые слова: экстремальные осадки, коэффициент хвостовой зависимости, копулы.

УДК 519.25

Разработка эффективного алгоритма краткосрочного прогнозирования электропотребления с использованием метода ансамбля

Е. Ю. Щетинин*, М. В. Бережков†, П. Г. Любин‡

* *Всероссийский научно-исследовательский институт
по проблемам гражданской обороны и чрезвычайных ситуаций
МЧС России (федеральный центр науки и высоких технологий)
ул. Давыдовская, д. 7, Москва, Россия, 121352*

† *ФГБОУ ВО Московский государственный технологический университет
«СТАНКИН»
пер. Вадковский, д. 3а, Москва, Россия, 127055*

Email: lyubin.p@gmail.com, riviera-molto@mail.ru

Краткосрочное прогнозирование потребления электроэнергии является актуальной задачей во многих областях человеческой деятельности в виду специфики продукта: нельзя накопить и хранить энергию впрок. Во-первых, предприятиям-участникам оптового рынка электроэнергии необходимо заранее подавать заявки с плановым потреблением, а энергогенерирующим предприятиям необходимо планировать мощности. Во-вторых, данный показатель может использоваться в качестве одного из признаков при построении других моделей. При этом потребление электрической энергии каким-либо объектом является временным рядом, так как представляет собой мгновенные значения потребляемой мощности замеры в различные моменты времени с определенной периодичностью. В данной работе продемонстрирован простой и эффективный метод краткосрочного прогнозирования электропотребления. Подход основывается на методе ансамбля базовых моделей (RPART — Recursive PARTitioning, CTREE — Conditional Inference Trees [5]) и имеет хороший уровень прогнозирования, который сопоставим с более сложными в использовании алгоритмами. Метод ансамбля представляет собой алгоритм комбинации набора обученных моделей с целью повышения точности прогноза, но стараясь избежать переобучения. Существует несколько методов ансамбля, которые имеют свои недостатки и преимущества. В данной работе мы использовали метод бэггинга (Bagging - Bootstrap aggregating [6]), который помог улучшить прогностическую силу отдельных базовых моделей.

Ключевые слова: краткосрочное прогнозирование, метод ансамбля, RPART, CTREE, случайные деревья, электропотребление, бэггинг.

1. Введение

К наиболее распространенным методам прогнозирования временных рядов относятся:

- прогнозная экстраполяция
- экспертные (интуитивные) методы прогнозирования
- корреляционный и регрессионный анализы
- прогнозирование на базе ARIMA моделей
- адаптивные методы прогнозирования
- прогнозирование с использованием искусственных нейронных сетей
- прогнозирование с использованием гибридных сетей

Перечисленные методы могут применяться для прогнозирования электропотребления и обладают присущими им достоинствами и недостатками. В современных работах чаще остальных описываются решения данной задачи с применением искусственных нейронных сетей, к недостаткам которых можно отнести сложность настройки и сложность интерпретации. В данной работе мы используем подход, в котором используется ансамбль моделей.

На рис. 1 изображена динамика почасового потребления электроэнергии в России за 3 недели 2017 года: с 13 июня по 3 июля.

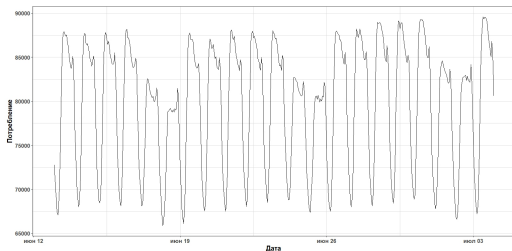


Рис. 1. Исходные данные

2. Метод

К наиболее распространенным методам ансамбля относятся:

- простое голосование (Simple Voting);
- взвешенное голосование (Weighted Voting);
- смесь экспертов (Mixture of Experts);
- бустинг (Boosting);
- бэггинг (Bagging — Bootstrap aggregating [6]).

В своей работе мы использовали бэггинг, который был предложен Л. Брейманом в 1996 году. Суть метода заключается в формировании различных обучающих подвыборок случайным выбором с возвращениями — некоторые объекты попадают в подвыборку несколько раз, некоторые ни разу. Базовые алгоритмы, обученные по подвыборкам, объединяются в композицию с помощью простого голосования. Достоинствами бэггинга являются: во-первых, возможность использования различных базовых алгоритмов, ошибки которых могут быть взаимно компенсированы при голосовании; во-вторых, некоторые обучающие подвыборки могут не содержать объекты-выбросы и алгоритм, построенный по этим подвыборкам, может оказаться точнее алгоритма, построенного по полной выборке. В данной работе в качестве базовых алгоритмов используются RPART и СТREE. Результат прогнозирования потребления электроэнергии на одни сутки вперед (4 июля 2017 года) приведен на рис. 2.

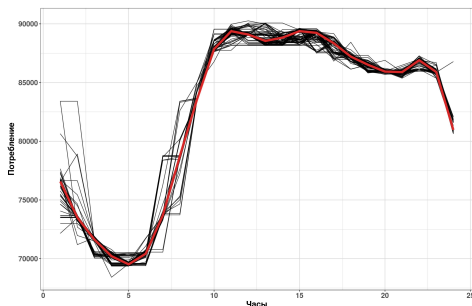


Рис. 2. Прогноз с использованием бэггинга

3. Заключение

В данной работе продемонстрирован простой и эффективный метод краткосрочного прогнозирования электропотребления, который может использоваться участниками рынка энергии при планировании генерации и при планировании закупок.

Литература

1. Щетинин Е. Ю. Эффективные компьютерные алгоритмы моделирования спотовых цен на электроэнергию // Научное обозрение. — 2016. — № 22. — С. 237–242.
2. Щетинин Е. Ю., Каплунов С. В., Марков П. Н. Моделирование спотовых цен на электроэнергию с использованием марковских процессов переключения режимов // Вестник РУДН. Серия: Математика. Информатика. Физика. — 2012. — № 3. — С. 61–68.
3. Щетинин Е. Ю., Любин П. Г. Робастный алгоритм построения сглаживающих сплайнов // Научное Обозрение. — 2015. — № 1. — С. 86–94.
4. Любин П. Г., Щетинин Е. Ю. Стохастические модели сглаживания и прогнозирования коэффициентов смертности // Научное Обозрение. — 2015. — № 18. — С. 147–155.
5. Breiman L., Friedman J. H., Olshen R. A., Stone C. J. Classification and Regression Trees. — Wadsworth, California.
6. Breiman L. Bagging Predictors // Machine Learning. — 1996. — Vol. 24. — Pp. 123–140.

UDC 519.25

Development of an Effective Algorithm for Short-Term Forecasting of Power Consumption Using Ensemble

E. Yu. Shchetinin*, M. V. Berezhkov[†], P. G. Lyubin[†]

** All-Russian Research Institute
for Civil Defense and Emergencies of the MESRF
(Science and High Technology Federal Center)*

Davydkovskaya str., 7, Moscow, 121352, Russian Federation

[†] Moscow State University of Technology “STANKIN”

Vadkovsky lane, 3a, Moscow, 127055, Russian Federation

Email: lyubin.p@gmail.com, riviera-molto@mail.ru

Short-term forecasting of electricity consumption is an actual task in many areas of human activity due to the specificity: consumers and power companies can't accumulate energy and can't store energy. At the first, consumers which participating in the wholesale electricity market must to submit a plan of future consumption, and energy-generating companies need to plan the output. Secondly, this indicator can be used as one of the features when fitting other models. At the same time, the consumption of electrical energy by any object is a time series, since it represents the instantaneous values of the consumed power measured at different times with a certain periodicity. In article we demonstrate a simple and effective method for short-term forecasting of power consumption. The approach is based on the method of base models ensemble (RPART — Recursive PARTitioning, CTREE — Conditional Inference Trees [5]) and has a good prediction level comparable with more complex algorithms. The ensemble method is an algorithm for combining a set of trained models to improve the accuracy of the forecast with trying to avoid overfitting. There are several ensemble methods that have their disadvantages and advantages. We used Bagging (Bootstrap aggregating [6]) which helped to improve the predictive power of particular base models.

Key words and phrases: short-term forecast, ensemble, RPART, CTREE, random trees, power consumption, bagging.

УДК 517.9

Методика отыскания алгебраических интегралов дифференциальных уравнений первого порядка

Юй Ин

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: yingy6165@gmail.com

В статье рассматривается практическое решение задачи Дебона: для заданного дифференциального уравнения $pdx + qdy = 0$, где p, q — многочлены из кольца $\mathbb{Q}[x, y]$, выяснить, имеет ли оно рациональный интеграл, и в случае утвердительного ответа предьявить этот интеграл. В основу предложенного подхода положен метод М.Н. Лагутинского. Теория и ее реализация протестированы на примерах из задачника А.Ф. Филиппова. Проведенные численные эксперименты свидетельствуют, что метода позволяет на практике без особых затрат ресурсов и времени идентифицировать наличие рационального интеграла, однако является весьма затратной как метод вычисления этого интеграла. Обсуждена проблема отыскания верхней грани для порядка интеграла и ее значение для решения дифференциальных уравнений на практике. В заключении даны рекомендации по оптимальному использованию метода М.Н. Лагутинского. Все вычисления выполнены в системе компьютерной алгебры Sage.

Ключевые слова: метод Лагутинского, задача Дебона, интегральные алгебраические кривые, алгебраические интегралы, sage.

1. Введение

При решении системы дифференциальных уравнений чрезвычайно важно найти все алгебраические интегралы. Эта задача, пусть и в других терминах, была предложена Дебоном Декарту еще в 1640-х годах, однако она до сих пор не решена: пользователь любого пакета, ищущего рациональный интеграл, должен задавать верхнюю границу для порядка искомого интеграла [2]. Обратимся к более простой, ограниченной задаче.

Задача. Для заданного дифференциального уравнения

$$p(x, y)dx + q(x, y)dy = 0, \quad p, q \in \mathbb{Q}[x, y], \quad (1)$$

и числа N , выяснить, имеет ли оно рациональный интеграл, порядок которого не превосходит N , и в случае утвердительного ответа предьявить этот интеграл.

Эта задача может быть решена методом неопределенных коэффициентов, однако на практике такое решение приводит к большим системам нелинейных уравнений на коэффициенты, решение которой приводит к большим вычислительным затруднениям. Накануне Первой мировой войны М.Н. Лагутинский разработал метод решения ограниченной задачи Дебона [1–3]. В настоящем докладе представлены результаты тестирования пакета Lagutinski [4] под Sage на дифференциальных уравнениях, взятых из задачника А.Ф. Филиппова [5].

2. Результаты апробации

На основе задач №№ 301–420 из задачника А.Ф. Филиппова был выделен список из 50 дифференциальных уравнений вида (1) 20 из которых имеют рациональные интегралы, а остальные интегрируются в элементарных функциях.

Прежде всего, следует заметить, что вычисление Δ_{55} в случайной точке позволило быстро выделить все 20 номеров, вероятно допускающих рациональные интегралы. При этом про оставшиеся 30 номеров можно с уверенностью сказать,

что их интегральные кривые или являются трансцендентными, или имеют порядок, больший 9-го.

Во-вторых, вычисление Δ_n в случайной точке позволило быстро подобрать n во всех номерах.

Вычисление определителя как многочлена относительно x, y заняло более часа в двух примерах: №№ 395 и 418. Вычисление интеграла заняло более часа в задаче № 361 и было остановлено. Во всех остальных 17 случаях вычисление интеграла было успешным во всех случаях, ответы совпадали с ответами из задачника после упрощений.

Рассмотрим проблемные номера.

Пример 1. Интегральные кривые уравнения № 395

$$(x^3 - 2xy^2)dx + 3x^2ydy = xdy - ydx$$

образуют семейство кривых

$$x(y^2 + x^2)^3 - 2/5y^5 - 4/3x^2y^3 - 2x^4y + cx^5 = 0$$

6-го порядка. Вычисление определителей в случайной точке однозначно указали на то, что вероятно $\Delta_{36} = 0$, однако вычислить этот определитель за час не удалось.

Пример 2. Интегральные кривые уравнения № 418

$$(3xy + x + y)ydx + (4xy + x + 2y)xdy = 0$$

образуют семейство кривых

$$6x^3y^4 + 2x^3y^3 + 3x^2y^4 = c$$

7-го порядка. Вычисление определителей в случайной точке однозначно указали на то, что вероятно $\Delta_{31} = 0$, однако вычислить этот определитель за час не удалось.

Пример 3. Интегральные кривые уравнения № 361

$$(2x^2y - 3y^2)y' = 6x^2 - 2xy + 1$$

образуют семейство

$$2x^3 - x^2y^2 + y^3 + x = c$$

4-го порядка. Мы вычислили $\Delta_{13} = 0, \Delta_{12} \neq 0$, поэтому точно существует рациональный интеграл порядка $N = 13$. Вычислить его не удалось.

Глядя в ответы, полученные, разумеется, другими методами, нетрудно понять, что предложенный подход позволил легко и быстро выявить все уравнения, допускающие алгебраические интегралы, но не для всех из них позволил найти явное выражение для интеграла. Наиболее проблемным, таким образом, является третий шаг.

3. Заключение

Проделанные численные эксперименты свидетельствуют, что метод Лагутинского позволяет на практике быстро и без особых затрат ресурсов и времени идентифицировать наличие рационального интеграла, однако является весьма затратным как метод вычисления этого интеграла. Следует заметить, что проблема отыскания границы для порядка интеграла, всегда обсуждаемая в теории, на практике оказалась не важной, в задачнике нет дифференциальных уравнений, интегральные кривые которых имели бы 10 порядок или более.

Вычисление определителей в случайных точках позволяет быстро выяснить, какие мономы должны фигурировать в выражении для интеграла, поэтому естественное направление оптимизации — предварительный подбор базиса по данным в

случайных точках. В частности, наиболее интересны те интегралы, которые можно записать как отношение малочленов. В частности, в пакете Lagutinski [4] можно делать перебор по всевозможным тройкам, составленных из первых N мономов кольца $\mathbb{Q}[x, y]$. Двигаясь в этом направлении, можно пытаться увеличить число элементов в кортеже, заменив затратное вычисление определителей Лагутинского вычислением их значения в случайной точке.

Благодарности

Автор признательна проф. Л. А. Севастьянову и доц. М. Д. Малых за постоянное внимание к ее работе.

Литература

1. Christopher C., Llibre J., and Pereira J. V. Multiplicity of Invariant Algebraic Curves in Polynomial Vector Fields // Pacific J. Math. — 2007. — Vol. 229 (1). — Pp. 63–117.
2. Chéze G. Computation of Darboux Polynomials and Rational First Integrals with Bounded Degree in Polynomial Time // Journal of Complexity. — 2011. — Vol. 27(2). — Pp. 246–262.
3. Малых М. Д. Об отыскании рациональных интегралов систем обыкновенных дифференциальных уравнений по методу М. Н. Лагутинского // Вестник НИЯУ МИФИ. — 2016. — Vol. 5(24). — Pp. 327–336.
4. Малых М. Д. О применении метода М. Н. Лагутинского к интегрированию дифференциальных уравнений 1-го порядка. Часть 1. Отыскание алгебраических интегралов // Вестник РУДН. Серия: Математика. Информатика. Физика. — 2017. — Т. 25, № 2. — С. 103–112.
5. Филиппов А. Ф. Сборник задач по дифференциальным уравнениям. — Ижевск: R&C, 2000.

UDC 517.9

The Method of Finding Algebraic Integral for First-order Differential Equations

Yu Ying

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
Miklukho-Maklaya str. 6, Moscow, 117198, Russian Federation*

Email: yingy6165@gmail.com

In the paper, the practical work of the Darboux problem is considered: for a given differential equation, we need to identify whether it is in the form of rational integral, and if the answer is true, we need to quadrature it. Our work is based on the method of M.N. Lagutinski. The theory and its realization are tested on the problems from Text-Book on Differential Equations by A.F. Filippov. The numerical experiments, which were carried out, show that the method makes it possible to identify the existence of the rational integral without taking much resource and time. However, using the method to calculate integrals is very time-consuming. The problem of finding the upper bound of the integral order and its value for solving differential equations practically are discussed. In the conclusion, how to optimally utilize the method of lagutinski is recommended. All calculations are executed in the computer algebra system Sage.

Key words and phrases: Lagutinski method, problem of Florimond de Beaune, integrated algebraic curves, algebraic integrals, sage.

УДК 004.4

Обзор методов решения задач мониторинга производительности и доступности ИТ структуры предприятия

А. В. Бабин

*Факультет повышения квалификации и переподготовки,
Московский технический университет связи и информатики,
ул. Авиамоторная, д. 8а. Москва, Россия, 111024*

Email: babinav@cbr.ru

Архитектура крупного современного предприятия чрезвычайно сложна. Использование сетей передачи данных в функционировании территориально распределенных подразделений с одной стороны облегчили работу их специалистов, с другой стороны усложнили обслуживание сетевого оборудования и серверной части. Широко распространенные облачные структуры, виртуальные вычислительные системы позволяют получать необходимые ресурсы для новых инфраструктур и приложений. Но географическое распределение и разнообразие используемых операционных систем глобально усложнили администрирование. В результате вопросы мониторинга ИТ инфраструктуры стали весьма актуальными.

В настоящее время любой программный продукт непрерывно видоизменяется в соответствии с ростом объема обрабатываемых данных и пользовательской аудиторией. Основная задача ИТ-специалистов — обеспечить бесперебойное функционирование непрерывно масштабируемых ИТ ресурсов. Чем сложнее решаемые предприятием задачи, тем сложнее ИТ инфраструктура, и тем глобальней фронт работ сотрудников, обеспечивающих работу системы.

Повысить доступность, производительность и эффективность работы сети, сократить время восстановления системы после аварии, вести профилактические работы, видеть состояние любого объекта сети позволяет внедрение систем мониторинга

Ключевые слова: система мониторинга, информационные технологии, архитектура сети, доступность и производительность.

1. Введение

Архитектура ИТ-структуры крупного современного предприятия чрезвычайно сложна. Это комплекс телекоммуникационных и аппаратных средств разных производителей, на базе системного ПО разных вендоров. От грамотного решения вопросов управления доступностью и производительностью ИТ ресурсов [1] зависит эффективность работы предприятия. Для повышения готовности приложений и систем, для минимизации среднего времени устранения неисправностей очень важно иметь средства идентификации и интеллектуальной диагностики, позволяющие точно выявлять и устранять основные причины возникающих проблем.

2. Основная часть

Для решения подобных задач широко используются системы мониторинга. С их помощью можно заблаговременно узнать о возможных сбоях и своевременно их устранять, получать детальную статистику нагрузки на объекты мониторинга, выявлять попытки взлома и несанкционированного доступа к информации, находить лимитирующие производительность компоненты [4]. Такая система может быть как обособленной, для выполнения конкретных задач на небольшом предприятии, так и являться составной частью системы управления информационной инфраструктурой предприятия. С помощью систем мониторинга решаются задачи самого широкого спектра: от систем контроля и управления доступом до комплексных систем централизованного представления информации, позволяющих получать

данные, осуществлять мониторинг и управлять связями между информационными системами и бизнес-процессами. Одной из задач, решаемых с помощью таких систем, является мониторинг производительности и доступности ИТ структуры предприятия, которую мы и рассмотрим более детально.

Системы мониторинга различны по способу реализации и внедрения, в зависимости от масштаба и архитектуры сети. Для крупных предприятий необходимо сначала определить объекты мониторинга (сетевое оборудование, серверы, приложения и т.д.), затем выбираются критичные показатели для каждого объекта. Правильная оценка показателей – залог эффективной работы системы, так как излишек данных перегрузит систему сообщениями об ошибках, а недостаток может привести к потере важных данных для анализа. После этого нужно определиться с архитектурой, выбрать продукт, решение. Затем система потребует настройки. Большие, сложные и дорогие решения требуют огромных трудозатрат по их разворачиванию и внедрению, но для крупного бизнеса это окупается. Есть варианты поменьше и попроще, они представляют собой так называемое «коробочное» решение, которое достаточно легко внедрить. Самое известное решение из недорогих — Microsoft SCOM. Есть ряд вариантов ПО с открытым исходным кодом [2] (например, Nagios, Centreon, Zabbix и т.д.) они бесплатны, но требовательны к уровню подготовки ИТ-специалистов, их обслуживающих.

Системы мониторинга производительности и доступности выполняют следующие функции:

- мониторинг неисправностей (аппаратной части: температура компонентов, параметры жестких дисков, целостность RAID дисковых массивов);
- мониторинг производительности (контроль постоянной и корректной работы компонентов ПО, контроль ключевых показателей производительности и нагрузки).

В самом общем виде мониторинг осуществляется следующими методами [3]:

1. Проверка физической доступности оборудования (например, проверка доступности сервера по протоколу ICMP (ping). Контролируется не только факт ответа, но и время задержки);
2. Проверка работоспособности запущенных служб и сервисов (TCP-подключение к соответствующему порту сервера, на котором должна быть запущена служба, и выполнение тестового запроса (например, аутентификации на почтовом сервере по протоколу SMTP или POP или запрос тестовой страницы от веб-сервера) Так же контролируется время задержки ответа).
3. Детальная проверка параметров функционирования сети: доступная и используемая память, свободное место на дисках, средняя нагрузка, количество процессов, количество пользователей и более специфические, такие как статус принтера у сервера печати.
4. Проверка параметров, специфичных для сервисов и служб данного конкретного окружения (наличие некоторых значений в таблицах БД, содержимое лог-файлов).

Наиболее важным фактором, влияющим на результаты работы системы, является грамотный анализ полученных значений мониторинга и правильная настройка пороговых значений.

3. Заключение

Резюмируя, следует отметить, что внедрение системы мониторинга производительности и доступности ИТ-инфраструктуры на предприятии позволяет исключить человеческий фактор при принятии решения в экстренной ситуации, снизить расходы на обслуживание сети, повысить отказоустойчивость ИТ – ресурсов, обеспечить оптимальное распределение и использование ресурсов сети, повысить эффективность работы ИТ службы предприятия.

Литература

1. Атчисон Ли. Масштабирование приложений. Выращивание сложных систем. –Спб.: Издательский дом «Питер», 2018
2. Список: системы мониторинга для Linux, бесплатные с открытым кодом. Электронный ресурс. <https://andreya.ru/operacionnaya-sistema-linux/spisok-sistemy-monitoringa-dlya-linux-besplatnye-i-otkrytum-kodom>
3. Эд Уилсон. Мониторинг и анализ сетей. Методы выявления неисправностей.- М.: Лори, 2012
4. Дон Джонс. «Создание унифицированной системы IT-мониторинга в вашем окружении». Электронный ресурс. <https://habrahabr.ru/post/173537/>

UDC 004.4

Overview of Technologies to Solve the Problems of Monitoring the Performance and Availability of the IT Structure of the Enterprise

A. V. Babin

*Faculty of advanced training and retraining
Moscow Technical University of Communication and Informatics,
8a Aviamotornaya str., Moscow, 111024, Russian Federation*

Email: babinav@cbr.ru

The architecture of a large modern enterprise is extremely complex. The use of data networks in the operation of geographically distributed units on the one hand facilitated the work of their specialists, on the other hand complicated the maintenance of network equipment and server part. Widespread cloud structures, virtual computing systems allow you to get the necessary resources for new infrastructures and applications. But the geographical distribution and variety of operating systems used globally complicated administration. As a result, IT infrastructure monitoring issues have become very relevant.

Currently, any software product is constantly changing in accordance with the growth in the volume of processed data and the user audience. The main task of IT-specialists is to ensure the uninterrupted operation of continuously scalable IT resources. The more complex the tasks solved by the company, the more complex IT infrastructure is, and the more global the front of the employees working to ensure the operation of the system.

Increase the availability, productivity and efficiency of the network, reduce the recovery time of the system after an accident, conduct preventive work, see the status of any network object allows the implementation of monitoring systems.

Key words and phrases: monitoring system, information technologies, network architecture, availability and productivity.

УДК 621.39

К анализу системы массового обслуживания фазового типа в частном случае

Е.В. Конькова

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Email: katiconkova1997@yandex.ru

Настоящая работа посвящена исследованию однолинейной системы массового обслуживания (СМО) фазового типа для небольших значений структурных параметров системы.

Ключевые слова: модель фазового типа, система массового обслуживания, система уравнений равновесия, рекуррентное уравнение.

1. Введение

Исследуется модель однолинейной системы массового обслуживания фазового типа с очередью $M|PH|1|г$. Символ «PH» обозначает распределение фазового типа [1]. В систему поступает пуассоновский поток заявок с интенсивностью λ . Прибывающая заявка, обнаруживающая полный буфер, теряется. Заявка начинает обслуживание на приборе на i -й фазе и может закончить обслуживание, обслужившись на данной фазе, либо продолжить обслуживание на фазе $(i+1)$. Таким образом, обслуживание разных заявок может состоять из разного числа фаз.

2. Исследование системы в частном случае

Для анализа однолинейной СМО фазового типа построим модель в частном случае с накопителем емкости $r = 2$ и количеством фаз на приборе $m = 3$ с заданными интенсивностями обслуживания, обозначаемыми как:

$$\vec{\mu} = (\mu_1, \mu_2, \mu_3).$$

Согласно вероятностной интерпретации PH-распределения компонента β_j вектора $\vec{\beta}$ есть вероятность начать обслуживание с фазы j . Параметр M_{ij} , интенсивность перехода заявки на приборе из фазы i в фазу j , задается с помощью следующей матрицы:

$$M = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{pmatrix}$$

Диаграмма интенсивностей переходов для рассмотренного случая изображена на рис. 1.

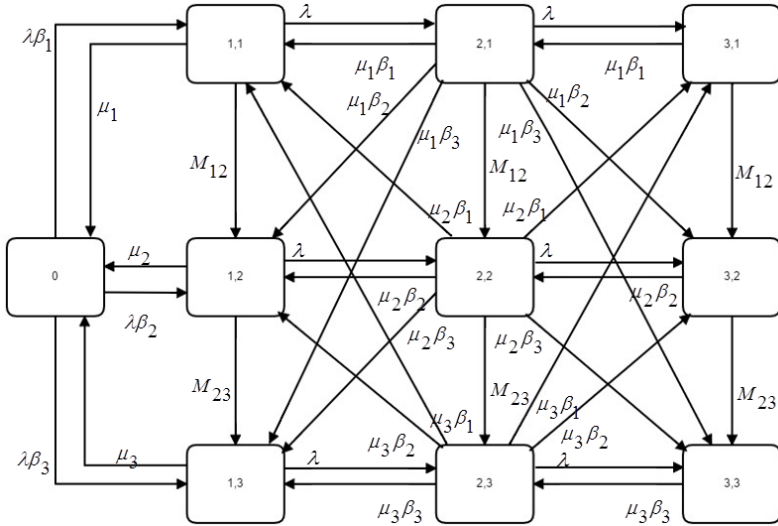


Рис. 1. Диаграмма интенсивностей переходов

Воспользовавшись принципом глобального баланса, построена СУР:

$$\left\{ \begin{array}{l} p_0 \lambda (\beta_1 + \beta_2 + \beta_3) = \mu_1 p_{11} + \mu_2 p_{12} + \mu_3 p_{13} \\ p_{11} (\lambda + \mu_1 + M_{12}) = \beta_1 (\lambda p_0 + \mu_1 p_{21} + \mu_2 p_{22} + \mu_3 p_{23}) \\ p_{12} (\lambda + \mu_2 + M_{23}) = \beta_2 (\lambda p_0 + \mu_1 p_{21} + \mu_2 p_{22} + \mu_3 p_{23}) + M_{12} p_{11} \\ p_{13} (\lambda + \mu_3) = \beta_3 (\lambda p_0 + \mu_1 p_{21} + \mu_2 p_{22} + \mu_3 p_{23}) + M_{23} p_{12} \\ p_{21} (\lambda + \mu_1 (\beta_1 + \beta_2 + \beta_3) + M_{12}) = \beta_1 (\mu_1 p_{31} + \mu_2 p_{32} + \mu_3 p_{33}) + \lambda p_{11} \\ p_{22} (\lambda + \mu_2 (\beta_1 + \beta_2 + \beta_3) + M_{23}) = \beta_2 (\mu_1 p_{31} + \mu_2 p_{32} + \mu_3 p_{33}) + \lambda p_{12} + M_{12} p_{21} \\ p_{23} (\lambda + \mu_3 (\beta_1 + \beta_2 + \beta_3)) = \beta_3 (\mu_1 p_{31} + \mu_2 p_{32} + \mu_3 p_{33}) + \lambda p_{13} + M_{23} p_{22} \\ p_{31} (M_{12} + \mu_1 (\beta_1 + \beta_2 + \beta_3)) = \lambda p_{21} \\ p_{32} (M_{23} + \mu_2 (\beta_1 + \beta_2 + \beta_3)) = \lambda p_{22} + M_{12} p_{31} \\ p_{33} \mu_3 (\beta_1 + \beta_2 + \beta_3) = \lambda p_{23} + M_{23} p_{32} \\ p_0 + \sum_{i=1, R=3, j=1, m=3} p_{ij} = 1 \end{array} \right.$$

или в матричной форме:

$$\begin{cases} -\lambda p_0 + \vec{p}_1^T \vec{\mu} = 0 \\ \vec{p}_1^T (-\lambda I + M) + \lambda \vec{\beta}^T p_0 + \vec{p}_1^T \vec{\mu} \vec{\beta}^T = \vec{0}^T \\ \vec{p}_2^T (-\lambda I + M) + \lambda \vec{p}_1^T + \vec{p}_3^T \vec{\mu} \vec{\beta}^T = \vec{0}^T \\ \vec{p}_3^T M + \lambda \vec{p}_2^T = \vec{0}^T \end{cases}$$

Данной системе соответствуют рекуррентные уравнения для стационарных вероятностей:

$$\begin{cases} \lambda p_0 = \mu_1 p_{11} + \mu_2 p_{12} + \mu_3 p_{13} \\ \lambda(p_{11} + p_{12} + p_{13}) = \mu_1 p_{21} + \mu_2 p_{22} + \mu_3 p_{23} \\ \lambda(p_{21} + p_{22} + p_{23}) = \mu_1 p_{31} + \mu_2 p_{32} + \mu_3 p_{33} \end{cases}$$

Заметим, что то же решение было получено в [2] в векторной форме с использованием матрично-геометрического метода:

$$\lambda p_k = \vec{p}_{k+1}^T \vec{\mu}, k = \overline{0, r}.$$

Решением СУР является p_{ij} - стационарная вероятность того, что в системе находится i заявок, а заявка на приборе обслуживается на фазе j . Стационарные вероятности состояний системы позволяют получить основные показатели эффективности СМО такие, как: среднее количество заявок в системе, среднее количество заявок в очереди, преобразование Лапласа-Стилтьеса (ПЛС) распределения времени пребывания, ожидания и обслуживания.

3. Заключение и задачи дальнейших исследований

В дальнейшем планируется провести численный анализ модели $M|PH|1|2$ и модели из [1] для случая небольшой размерности.

Литература

1. Bocharov P.P., Naumov V.A. "Matrix-geometric stationary distribution for the PH|PH|1|R queue" // HAL archives-ouvertes, May 1984, p. 18
2. Бочаров П.П., Печинкин А.В. Теория массового обслуживания: Учебник. – М.: Изд-во РУДН, 1995. – 529 с., ил.

UDC 621.39

To the Analysis of Phase-type Queuing System in the Special Case

E. V. Konkova

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya str., Moscow, 117198, Russian Federation*

Email: katikonkova1997@yandex.ru

This work is devoted to the study of a single-line queuing system of phase-type for small structural parameters of the system.

Key words and phrases: phase-type model, queuing system, balance equation, recurrent equation.

Авторский указатель

- А**
- Абаев П. О. 87
 Авзалов Р. Л. 273
 Агеев К. А. 17
 Адам Г. 245
 Аду К. И. 20
 Аль-Натор М. С. 279
 Аль-Натор С. В. 279
 Амирханов И. В. 282
 Андреев С. Д. 90
 Анисимова А. А. 54
 Апреутсей А.-М. Ю. 285
 Аронов Д. А. 165
 Ахметханова Д. М. 288
- Б**
- Бабин А. В. 419
 Багаева Н. В. 23
 Байло Н. В. 291
 Балашов Н. А. 262
 Баранов А. В. 262
 Баскаков А. Е. 98
 Батгэрэл Б. 305
 Бахтин А. А. 93
 Башапин М. В. 248
 Бегипшев В. О. 137
 Белов М. П. 336
 Беляков Д. В. 245
 Березков М. В. 413
 Бесчастный В. А. 143, 154
 Богданова Е. В. 26
 Богданова Е. Ю. 96
 Боровских Д. А. 294
 Бочкарёв В. В. 297
 Будочкина С. А. 300
 Буй Д. Т. 69
 Бутенко Ю. А. 273
 Буурулдай А. Э. 302, 404
- В**
- Валя М. 245
 Васильева Д. Г. 387
 Васильюк И. П. 65
 Веллева Т. Р. 308, 311
 Волков А. С. 98, 101
 Волков С. С. 159
 Вольнец А. С. 104
- Г**
- Гаджизаде Н. К. 44
 Гайдамака А. А. 107
 Гайдамака Ю. В. 121
 Галпина О. С. 134
- Гарибян А. А. 104
 Геворкян М. Н. 314, 318, 361
 Герценбергер К. В. 273
 Гоголева С. Ю. 321
 Гольдштейн Б. С. 110
 Гольская А. А. 29
 Гончаров Л. М. 210
 Гончаров П. В. 245
 Горшенин А. К. 32
 Гостев И. М. 318
 Грачев Д. А. 324
 Григорьева Т. В. 162
 Гудкова И. А. 115, 134, 143, 154
 Гумбатов Р. Т. 38
- Д**
- Дараселия А. В. 118
 Демидова А. В. 314, 318, 329
 Демидова Т. С. 314
 Денисенко А. П. 333
 Дзантиев И. Л. 29
 Диваков Д. В. 336, 339, 401
 Дилие Гбавле Люк-Дональд . 294, 326
 Димова С. Н. 305
 Добрынин В. Н. 252
 Древицкий А. С. 339
 Дружинина О. В. 329
 Дубнов Ю. А. 257
 Дудин А. Н. 35
 Дудин С. А. 35
 Дудина О. С. 35
- Е**
- Егоров А. А. 342
 Емельянова Е. Ю. 324, 326
- Ж**
- Жданков А. Н. 121
 Жуков В. В. 162, 165, 174
 Жукова Л. В. 345
- З**
- Завозина А. В. 308, 311
 Зайнетдинова Л. Г. 321
 Зарядов И. С. 26, 60, 81, 288
 Земляная Е. В. 248
 Зотова П. А. 345, 381
 Зрелов П. В. 245
 Зубрихина М. О. 168
 Зуев М. И. 260
- И**
- Ибрагимов Б. Г. 38, 44
 Ибрагимов Р. Ф. 38

Иванова Д. В.	49
Исаев А. М.	44
Исаев Ю. Д.	213

К

Казандаева А. П.	381
Карачанская Е. В.	124
Карнаухова Е. А.	49
Касимов Ю. Ф.	348
Катрич А. В.	171
Керимов В. Р.	44
Кименчежи В. В.	127
Ковальчуков Р. Н.	52
Козловский А. Н.	174
Козырев Д. В.	75, 127
Коновалов М. Г.	351
Конькова Е. В.	422
Коняев Е. Ю.	177
Копылов А.	222
Кореньков В. В.	245
Королькова А. В.	308, 311, 318, 361
Котюков М. М.	339
Кочанов А. А.	396
Крескин А. Д.	180
Крупко О. С.	52
Кряев А. В.	297, 355
Кузив Я. Ю.	358
Кузнецов Е. А.	183
Кулябов Д. С.	318, 361
Кутбитдинов С. Ш.	131
Кутовский Н. А.	262

Л

Лазарев А. В.	101
Лесик П. А.	364
Лижненко Н. М.	366
Лохмотко В. В.	131
Лукин А. В.	186
Лукьянов К. В.	248
Любин П. Г.	413

М

Мажитова Е. М.	262
Макеева Е. Д.	134
Мардашев А. М.	210
Маркова Е. В.	20, 29, 49, 134
Масина О. Н.	329
Матвеев М. А.	245, 260
Матвеева Е. А.	189
Матюхин В. В.	355
Матюшенко С. И.	291, 381
Мацкевич И. А.	140
Мачнев Е. А.	137
Милованова Т. А.	26, 60, 81
Миловидова А. А.	252
Мингажитдинова Э. Ф.	333
Мияйлович Н.	329
Моисеева Е. В.	165

Мокров Е. В.	146
Молодченков А. И. 168, 192, 201, 207, 216	
Молчанов Д. А.	52
Морозова Д. А.	366
Москалева Ф. А.	52
Муратчаев С. С.	98
Мысина Ю. А.	368

Н

Назаров А. А.	54
Никитина В. В.	57
Никольский И. М.	266
Новикова Д. С.	195

О

Орбелян Т. С.	60
Острикова Д. Ю.	143, 154

П

Павлюкова А. В.	198
Пальчевский А. И.	201
Паршина К. С.	204
Перетьяко О. Д.	207
Пермякова Ю. С.	371
Петров В. А.	294, 333
Погосян Ш. Г.	63
Подгайный Д. В.	245, 260
Половов М. П.	143
Поляков Н. А.	134
Пономарева Л. А.	65
Попков А. Ю.	374
Попова Е. В.	324, 326
Пузынин И. В.	305
Пузынина Т. П.	305
Пылаева А. Н.	377
Пяткина Д. А.	291, 345, 381

Р

Разгоняев В. А.	146
Разумчик Р. В.	351
Рассахан Н. Д.	384
Рахмонова А. Р.	269
Романов А. А.	364
Романовская Ю. А.	115
Ромашкова О. Н.	65, 150
Рудинская С. Р.	131

С

Савин А. С.	294, 333
Савич В. Н.	154
Садовников В. К.	294
Салпагаров С. И.	210, 213
Самойлов В. Е.	150
Самуйлов А. К.	52
Самуйлов К. Е.	146
Саркар Н. Р.	282
Севастьянов А. Л.	336
Севастьянов Л. А.	387

Семенов Р. Н.	262
Семёнова О. В.	69
Семушина С. Г.	165
Сингх Л.	216
Синицын И. Э.	72
Скрынник К. В.	219
Слива Д. Е.	355
Смирнов Д. С.	297
Соболев А. А.	314
Соколов И. А.	252
Соколов Н. А.	110
Солодков А. В.	98
Соппин Э. С.	17
Соседова Н. И.	124
Соченков И. В.	235
Стрельцова О. И.	245, 260

Т

Тханг Фам К.	222
Тхао Чан Т. Т.	222
Талагаев Ю. В.	390
Тевелева Е. А.	393
Терлецкий А. Я.	393
Тимербаев М. И.	348
Торосян Ш. Г.	260
Третьяков Н. П.	393, 396
Тухлиев З. К.	305

У

Уанкпо Г. Ж. К.	75
----------------------	----

Ф

Фомин М. Б.	183, 229
------------------	----------

Х

Халина В. А.	78
Хамидуллина К. И.	399
Харин П. А.	134
Хачумов В. М.	226

Хачумов М. В.	226
Хохлов А. А.	294, 324, 326, 333
Храбров Р. Н.	229
Христов И. Г.	305
Христова Р. Д.	305
Хуракай Д. М.	57

Ц

Царев А. С.	87
Цурлуков В. В.	81

Ч

Чеботов А. И.	273
Чернышев Я. М.	273
Чистова Е. В.	232
Чистотина Д. А.	364
Чупритский В. К.	401
Чухно Н. В.	107
Чухно О. В.	107

Ш

Шарафутдинов Р. Б.	297
Шарипов З. А.	305
Шебанова К. В.	399
Шорохов С. Г.	404
Шунин Е. В.	407

Щ

Щетинин Е. Ю.	384, 410, 413
--------------------	---------------

Ю

Юй Ин	416
-------------	-----

Я

Ядринцев В. В.	235
Янко Ю. Д.	238
Ячимович М.	329

Author index

A

Abaev P. O.	87
Adam Gh.	245
Adou K. Y.	20
Ageev K. A.	17
Ahmethanova D. M.	288
Al-Nator M. S.	279
Al-Nator S. V.	279
Amirkhanov I. V.	282
Andreev S. D.	90
Anisimova A. A.	54
Apreutesey A.-M. Y.	285
Aronov D. A.	165
Avzalov R. L.	273

B

Babin A. V.	419
Bagaeva N. V.	23
Bakhtin A. A.	93
Balashov N. A.	262
Baranov A. V.	262
Bashashin M. V.	248
Baskakov A. E.	98
Batgerel B.	305
Baylo N. V.	291
Begishev V. O.	137
Belov M. P.	336
Belyakov D. V.	245
Berezhkov M. V.	413
Beschastnyi V. A.	143, 154
Bochkarev V. V.	297
Bogdanova E. V.	26
Bogdanova E. Y.	96
Borovskikh D. A.	294
Budochkina S. A.	300
Bui D. T.	69
Butenko Y. A.	273
Buuruldai A. E.	302, 404

C

Chebotov A. I.	273
Chernyshev Y. M.	273
Chistotina D. A.	364
Chistova E. V.	232
Chukhno N. V.	107
Chukhno O. V.	107
Chupritskiy V. K.	401

D

Daraseliya A. V.	118
Demidova A. V.	314, 318, 329
Demidova T. S.	314
Denisenko A. P.	333

Dilie Gbavle Luk-Donald	294, 326
Dimova S. N.	305
Divakov D. V.	336, 339, 401
Dobrynin V. N.	252
Drevitskiy A. S.	339
Druzhinina O. V.	329
Dubnov Yu. A.	257
Dudin A. N.	35
Dudin S. A.	35
Dudina O. S.	35
Dzantiev I. L.	29

E

Egorov A. A.	342
Emelyanova E. U.	324, 326

F

Fomin M. B.	183, 229
------------------	----------

G

Gaidamaka A. A.	107
Gaidamaka Yu. V.	121
Galinina O. S.	134
Gariyban A.A.	104
Gertsenberger K. V.	273
Gevorkyan M. N.	314, 318, 361
Gogoleva S. Y.	321
Goldstein B. S.	110
Golskaia A. A.	29
Goncharov L. M.	210
Goncharov P. V.	245
Gorshenin A. K.	32
Gostev I. M.	318
Grachev D. A.	324
Grigoryeva T. V.	162
Gudkova I. A.	115, 134, 143, 154

H

Hajzade N. K.	44
Harin P. A.	134
Houankpo H. G. K.	75
Hristov I. G.	305
Hristova R. D.	305
Humbatov R. T.	38

I

Ibrahimov B. G.	38, 44
Ibrahimov R. F.	38
Isaev Yu. D.	213
Isayev A. M.	44
Ivanova D. V.	49

J

Jacimovic M.	329
-------------------	-----

K

Karachanskaya E. V.	124
Karimov V. R.	44
Karnauhova E. A.	49
Kasimov Y. F.	348
Katrish A. V.	171
Kazandaeva A. P.	381
Khachumov M. V.	226
Khachumov V. M.	226
Khalina V. A.	78
Khamidulina K. I.	399
Khokhlov A. A.	294, 324, 326, 333
Khrabrov R. N.	229
Khurakay D. M.	57
Kimenezhi V. V.	127
Kochanov A. A.	396
Konkova E. V.	422
Konovalov M. G.	351
Konyaev E. Yu.	177
Kopylov A.	222
Korenkov V. V.	245
Korolkova A. V.	308, 311, 318, 361
Kotlyukov M. M.	339
Kovalchukov R. N.	52
Kozlovskii A. N.	174
Kozyrev D. V.	75, 127
Kreskin A. D.	180
Krupko A. S.	52
Kryanev A. V.	297, 355
Kulyabov D. S.	318, 361
Kutbitdinov S. Sh.	131
Kutovskiy N. A.	262
Kuziv Ia. lu.	358
Kuznetsov E. A.	183

L

Lazarev A. V.	101
Lesik P. A.	364
Lizhnenko N. M.	366
Lokhmotko V. V.	131
Lukin A. V.	186
Lukyanov K. V.	248
Lyubin P. G.	413

M

Machnev E. A.	137
Makeeva E. D.	134
Mardashev A. M.	210
Markova E. V.	20, 29, 49, 134
Masina O. N.	329
Matokhin V. V.	355
Matskevich I. A.	140
Matveev M. A.	260
Matveeva E. A.	189
Matveyev M. A.	245
Matyushenko S. I.	291, 381
Mazhitova E. M.	262
Mijajlovic N.	329

Milovanova T. A.	26, 60, 81
Milovidova A. A.	252
Mingazhitdinova E. F.	333
Moiseeva E. V.	165
Mokrov E. V.	146
Molodchenkov A. I.	168, 192, 201, 207, 216
Moltchanov D. A.	52
Morozova D. A.	366
Moskaleva F. A.	52
Muratchaev S. S.	98
Mysina J. A.	368

N

Nazarov A. A.	54
Nikitina V. V.	57
Nikolsky I. M.	266
Novikova D. S.	195

O

Orbelyan T. S.	60
Ostrikova D. Yu.	143, 154

P

Palchevskii A. I.	201
Parshina K. S.	204
Pavlyukova A. V.	198
Peretyatko O. D.	207
Permyakova J. S.	371
Petrov V. A.	294, 333
Podgainy D. V.	245, 260
Pogosian Sh. G.	63
Polovov M. P.	143
Polyakov N. A.	134
Ponomareva L. A.	65
Popkov A. Y.	374
Popova E. V.	324, 326
Puzynin I. V.	305
Puzynina T. P.	305
Pyatkina D. A.	291, 345, 381
Pylaeva A. N.	377

R

Rahmonova A. R.	269
Rassakhan N. D.	384
Razgonyaev V. A.	146
Razumchik R. V.	351
Romanov A. A.	364
Romanovskaya J. A.	115
Romashkova O. N.	65, 150
Rudinskaya S. R.	131

S

Sadovnikov V. K.	294
Salpagarov S. I.	210, 213
Samouylov K. E.	146
Samoylov V. E.	150
Samuilov A. K.	52
Sarker N. R.	282

Savich V. N.	154
Savin A. S.	294, 333
Semenov R. N.	262
Semenova O. V.	69
Semushina S. G.	165
Sevastianov A. L.	336
Sevastianov L. A.	387
Sharafutdinov R. B.	297
Sharipov Z. A.	305
Shchetinin E. Yu.	384, 413
Shebanova K. V.	399
Shorokhov S. G.	404
Shunin E. V.	407
Singh L.	216
Simitsyn I. E.	72
Skrinnik K. V.	219
Sliva D. E.	355
Smirnov D. S.	297
Sobolev A. A.	314
Sochenkov I. V.	235
Sokolov I. A.	252
Sokolov N. A.	110
Solodkov A. V.	98
Sopin E. S.	17
Sosedova N. I.	124
Streltsova O. I.	245, 260

T

Thang Pham Cong	222
Thao Tran Thi Thu	222
Talagaev Y. V.	390
Terletski A. Ya.	393
Teveleva E. A.	393
Timerbaev M. I.	348

Torosyan Sh. G.	260
Tretyakov N. P.	393, 396
Tsarev A. S.	87
Tsurlukov V. V.	81
Tukhliev Z. K.	305

V

Vala M.	245
Vasilyeva D. G.	387
Vasilyuk I. P.	65
Velieva T. R.	308, 311
Volkov A. S.	98
Volkov S. S.	159
Volkov V. S.	101
Volynets A.S.	104

Y

Yadrintsev V. V.	235
Yanko J. D.	238
Yu Ying	416

Z

Zainetdinova L. G.	321
Zaryadov I. S.	26, 60, 81, 288
Zavozina A. V.	308, 311
Zemlyanaya E. V.	248
Zhdankov A. N.	121
Zhukov V. V.	162, 165, 174
Zhukova L. V.	345
Zotova P. A.	345, 381
Zrelov P. V.	245
Zubrikhina M. O.	168
Zuev M. I.	260

Научное издание

**ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫЕ
ТЕХНОЛОГИИ И МАТЕМАТИЧЕСКОЕ
МОДЕЛИРОВАНИЕ
ВЫСОКОТЕХНОЛОГИЧНЫХ СИСТЕМ**

Издание подготовлено в авторской редакции

Технический редактор *Н. А. Ясько*
Компьютерная вёрстка *А. В. Королькова, Д. С. Кулябов*
Дизайн обложки *Ю. Н. Ефремова*

Подписано в печать 10.04.2018 г. Формат 60×84/16.
Бумага офсетная. Печать офсетная. Гарнитура Таймс
Усл. печ. л. 24,88. Тираж 150 экз. Заказ 407.

Российский университет дружбы народов
115419, ГСП-1, г. Москва, ул. Орджоникидзе, д. 3

Типография РУДН
115419, ГСП-1, г. Москва, ул. Орджоникидзе, д. 3, тел. +7(495)952-04-41