Правительство Российской Федерации
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»
(НИУ ВШЭ)

УДК 33:005
Рег. № НИОКТР АААА-А19-119061190061-1
Рег. № ИКРБС

УТВЕРЖДАЮ
Проректор НИУ ВШЭ
канд. экон. наук
_____М.М. Юдкевич
«___» _____2019 г.

ОТЧЕТ
О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ

ИССЛЕДОВАНИЕ МОДЕЛЕЙ ПРИНЯТИЯ РЕШЕНИЙ И АНАЛИЗА СЛОЖНО СТРУКТУРИРОВАННЫХ ДАННЫХ
(заключительный)

Руководители НИР:
зав. МНУЛ анализа и выбора решений,
д-р. техн. наук _____Ф.Т. Алескеров

гл. научн. сотр. МНУЛ анализа и выбора
решений, PhD _____Э.С. Маскин

Москва 2019

СПИСОК ИСПОЛНИТЕЛЕЙ

Соруководители темы:

Зав. лабораторией, д.т.н., рук. деп., проф.	_____	Ф.Т. Алескеров	(введение, заключение)
Гл. научн. сотр., PhD	_____	Э. Маскин	(пункт 1.3.1)

Исполнители:

Ст.-исследователь		А.Р. Агишева	(пункт 1.5.1)
Профессор, д.т.н., проф.	_____	А.С. Беленький	(пункты 3.2.2 – 3.2.4)
Ст.-исследователь		Е.А. Борсук	(список исп. источн.)
Ст.-исследователь	_____	Ю.А. Веселова	(пункт 1.1.3)
Ст.-исследователь	_____	А.С. Власов	(пункт 2.2.2)
Ст.-исследователь	_____	И.Е. Гавриленкова	(пункт 1.4.1)
Ст.-исследователь	_____	Ю.Э. Гурова	(пункт 2.1.4)
Ст.-исследователь, аспирант	_____	К.А. Деева	(пункты 4.2.2 – 4.2.5)
Ст.-исследователь, аспирант	_____	С.С. Демин	(пункты 6.1.3, 6.1.4)
Доцент, PhD, доцент	_____	Э. Доган	(реферат)
Доцент, к.ф.-м.н., доцент	_____	Л.Г. Егорова	(пункты 2.3.4, 2.3.5)
Ст.-исследователь, аспирант	_____	М.Д. Ермолова	(пункты 5.1.1 – 5.1.3)
Ст.-исследователь		Е.С. Замятина	(пункт 4.2.1)
Ст.-исследователь	_____	А.А. Иванов	(пункт 1.1.2)
Доцент, к.э.н., доцент	_____	Д.С. Карабекян	(пункты 1.1.1, 1.1.4)

Доцент, к.э.н., доцент	_____	А.В. Карпов	(пункты 1.4.2, 1.4.3, 1.5.2, 1.5.3)
Ст. науч. сотр., д.т.н.	_____	Н.А. Коргин	(подраздел 3.1)
Вед. науч. сотр., д.ф.-м.н., проф.	_____	А.А. Лазарев	(пункт 3.2.1)
Профессор, д.ф.-м.н., проф.	_____	А.Е. Лепский	(пункты 4.1.1 – 4.1.4)
Ст.-исследователь, аспирант	_____	Т.А. Ломоносов	(пункты 6.1.1, 6.1.2)
Ст.-исследователь, аспирант	_____	Н.Г. Мещерякова	(пункт 2.1.3)
Профессор, д.т.н., проф.	_____	Б.Г. Миркин	(пункты 2.2.1, 2.2.4)
Ст. науч. сотр., к.т.н., доцент	_____	А.Л. Мячин	(пункты 2.3.1 - 2.3.3)
Ст.-исследователь		Ю.В. Неваленная	(пункты 4.1.5 - 4.1.6)
Ст. науч. сотр., к.э.н., доцент	_____	Г.И. Пеникас	(пункты 5.1.4, 5.1.5)
Проф.-исслед., д.т.н., проф.	_____	В.В. Подиновский	(подраздел 1.2)
Ст.-исследователь	_____	А.Н. Резяпова	(пункт 2.1.5)
Ст.-исследователь, студент	_____	А.А. Роман	(пункт 2.1.6)
Эксперт, к.т.н., доцент	_____	А.А. Рубчинский	(подраздел 5.2)
Доцент, к.ф.-м.н., доц.	_____	А.Н. Субочев	(пункты 1.3.2 – 1.3.7)

Ст.-исследователь	_____	М.А. Сурков	(пункты 5.1.6, 5.1.7)
Ст.-исследователь	_____	Д.Н. Тверской	(подраздел 6.2)
Ст.-исследователь	_____	Д.С. Фролов	(пункт 2.2.3)
Мл. науч. сотр., аспирант	_____	С.В. Швыдун	(пункт 2.1.2)
Ст. науч. сотр., к.т.н.	_____	В.И. Якуба	(пункты 2.1.1, 2.1.7)
Нормоконтроль	_____	А.А. Павлова	

РЕФЕРАТ

Отчет 294 с., 1 кн., 32 рис., 17 табл., 320 источника.

ТЕОРИЯ ПРИНЯТИЯ РЕШЕНИЙ, МЕТОДЫ ОПТИМИЗАЦИИ, ТЕОРИЯ КОЛЛЕКТИВНОГО ВЫБОРА, СЕТЕВОЙ И КЛАСТЕРНЫЙ АНАЛИЗ ДАННЫХ, ТЕОРИЯ ФУНКЦИЙ ДОВЕРИЯ, АНАЛИЗ ПАТТЕРНОВ

Объектами исследования проекта были методы многокритериального принятия решений, теории коллективного выбора, кластерного и сетевого анализа данных, анализа паттернов, теории функций доверия, модели в социально-экономических и банковских системах, динамические модели распространения и модели специализаций в многоклеточных системах.

Целью проекта было исследование моделей принятия решений и анализа данных. В направлении исследований моделей принятия решений речь идет о моделях конкурентной кооперации, моделях оптимального размещения, моделях теории расписаний и моделях специализации. Были продолжены теоретические исследования отношений предпочтения в теории коллективного выбора, ранжирований вариантов и правил выбора оптимальных альтернатив, решений при нечисловом оценивании предпочтений и неопределенности на основе теории важности критериев. В направлении анализа данных основное внимание было уделено анализу сетевых структур, анализу и интерпретации данных сложной структуры. Было продолжено исследования функционирования участников финансового рынка. Приложения теоретических исследований были связаны с решениями задач в социально-экономической, политической, финансовой и банковской сферах.

Методологической базой проекта были: современные методы теории принятия решений, методы оптимизации, теория игр и экономических механизмов, теория графов, теории многокритериального и коллективного выбора, интеллектуальные методы анализа данных, теории вероятностей и неточных вероятностей, теории функций доверия.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ 10
1 Исследование моделей выбора и агрегирования 19
1.1 Построение и применение теоретических моделей для расчета степени манипулируемости правил агрегирования предпочтений 19
1.1.1 Введение 19
1.1.2 Манипулирование при контругрозе 20
1.1.3 Безопасное коалиционное манипулирование 26
1.1.4 Некоторые выводы по исследованию манипулирования 30
1.2 Разработка методов анализа решений при нечисловом оценивании предпочтений и неопределенности на основе теории важности критериев 31
1.2.1 Введение 31
1.2.2 Математическая модель проблемной ситуации 32
1.2.3 Отношения вероятности 35
1.2.4 Построение отношений предпочтения и безразличия 37
1.2.5 Случай порядковой вероятности 41
1.2.6 Отношения предпочтения и безразличия при известном отношении ЛПР к риску 43
1.2.7 Отношения предпочтения и безразличия при шкале оценок первой порядковой метрики 46
1.2.8 Отношения предпочтения и безразличия при информации о сравнительной степени превосходства в вероятности 48
1.2.9 Некоторые выводы по исследованию 49
1.3 Аксиоматический анализ способов выбора оптимальных вариантов 50
1.3.1 Выбор, предпочтения, полезность 50
1.3.2 Аксиомы коллективного выбора 55
1.3.3 Теоремы о невозможности выбора 57
1.3.4 Отказ от аксиомы Транзитивности 59
1.3.5 Отказ от аксиомы Неограниченности области определения 62
1.3.6 Отказ от аксиомы Независимости от посторонних полезностей 63
1.3.7 Отказ от аксиомы Ординальной несравнимости 67
1.4 Исследование свойств профилей предпочтений 68
1.4.1 Введение 68

1.4.2 Модель и основные результаты 70
1.4.3 Некоторые выводы 72
1.5 Исследование условий согласованной единственности стабильного паросочетания 73
1.5.1 Введение 73
1.5.2 Задача о марьяже 74
1.5.3 Некоторые выводы 76
2 Разработка, исследование и применение методов анализа данных 78
2.1 Исследования по сетевому анализу данных 78
2.1.1 Введение 78
2.1.2 Оценка влияния в мультиплексных сетях 83
2.1.3 Меры влияния связей в сетевых структурах 89
2.1.4 Критические группы в сетях и меры центральности: проверка выполнимости аксиом. 94
2.1.5 Сетевой анализ международной миграции 96
2.1.6 Сетевой анализ международной торговли 98
2.1.7 Некоторые выводы 100
2.2 Исследования по анализу текстов и кластерному анализу 102
2.2.1 Введение и постановка задачи оптимального обобщения нечетких множеств 102
2.2.2 Экономный подъем нечеткого тематического кластера в таксономии: модель и метод 105
2.2.3 Структурирование коллекции научных публикаций с помощью таксономии предметной области 111
2.2.4 Некоторые выводы 120
2.3 Разработка методов анализа паттернов 122
2.3.1 Введение 122
2.3.2 Порядково-инвариантная паттерн-кластеризация 123
2.3.3 Примеры порядково-инвариантных паттерн-кластеров с взаимно симметричной структурой кусочно-линейных функций 124
2.3.4 Выделение порядково-инвариантных паттерн-кластеров с взаимно симметричной структурой данных. 134
2.3.5 Некоторые выводы 136
3 Исследования организационных моделей и механизмов 137
3.1 Исследования конкурентной кооперации в Арктике 137
3.2 Исследование моделей по оптимальному размещению объектов и привлечению частных инвестиций 142
3.2.1 Введение 142

3.2.2 Разработка и исследование модифицированной модели размещения грузового транспортного узла 147
3.2.3 Выявление системных проблем размещения объектов и привлечения частных инвестиций 165
3.2.4 О формализации и математическом обосновании утверждений 172
4. Исследования моделей анализа политических систем 176
4.1 Анализ значимости позиций партий для голосования по ним методами теории монотонных мер 176
4.1.1 Введение и постановка задачи исследования 176
4.1.2 Некоторые сведения из теории неаддитивных мер, интегралов и функций доверия 177
4.1.3 Постановка основных задач исследования 181
4.1.4 Описание данных 186
4.1.5 Результаты исследования 187
4.1.6 Некоторые выводы 189
4.2 Исследования поляризованности и влияния в политических системах 190
4.2.1 Введение 190
4.2.2 Политическое влияние в парламентах 192
4.2.3 Почему важно изучать поляризацию 195
4.2.4 Контекст изучения поляризованности 197
4.2.5 Некоторые выводы 200
5. Исследования моделей в финансовой и банковской сферах 201
5.1 Исследование влияния модельного риска на точность оценок величины риск-взвешенных активов 201
5.1.1 Общая постановка задачи 201
5.1.2 Обзор литературы 203
5.1.3 Описание методологии исследования 204
5.1.4 Причина недооценки кредитного риска 210
5.1.5 Пример количественной оценки недоучета кредитного риска 211
5.1.6 Калибровка надбавки 213
5.1.7 Основные выводы 216
5.2 Анализ структурной сложности данных по фондовому рынку S&P-500 217
5.2.1 Введение 217
5.2.2 Частотный алгоритм дихотомии 218
5.2.3 Сцепленность графа 220

5.2.4 Вычисление сцепленности для фондового рынка 222
5.2.5 Анализ сцепленности для фондового рынка 225
5.2.6 Некоторые выводы 227
6 Исследование моделей распространения и специализации 228
6.1 Исследование моделей распространения 228
6.1.1 Введение 228
6.1.2 Исследование решений одномерных уравнений газовой динамики 230
6.1.3 Задача оптимального расположения сенсоров обнаружения загрязнений на поверхности моря 231
6.1.4 Некоторые выводы 234
6.2 Исследование математических моделей возникновения специализации в многоклеточных системах 235
6.2.1 Введение 235
6.2.2 Основные результаты 239
6.2.3 Случай системы с элементами одного типа 241
6.2.4 Некоторые выводы 242
ЗАКЛЮЧЕНИЕ 244
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ 248
ПРИЛОЖЕНИЕ А Таблица к подразделу 5.2 278

ВВЕДЕНИЕ

В отчете представлены результаты работы в 2019 году сотрудников Международной научно-учебной лаборатории анализа и выбора решений (МЛАВР) НИУ ВШЭ над проектом «Исследование моделей принятия решений и анализа сложно структурированных данных».

В соответствии с техническим заданием были проведены исследования в следующих научных направлениях:

– исследование моделей выбора и агрегирования;

– разработка, исследование и применение методов анализа данных;

– исследование организационных моделей и механизмов;

– исследование моделей анализа политических систем;

– исследование моделей в финансовой и банковской сферах;

– исследование моделей распространения и специализации.

В данном проекте сотрудники МЛАВР продолжили исследования 2018 г. в ходе выполнения НИР «Современный контекст методов принятия решений и анализа данных: человеческий фактор, неопределенность, риски, сетевые модели, большие данные» [1]. В частности, в 2018 г. были получены следующие результаты:

– математически поставлены задачи об оценке окупаемости требуемых инвестиций в строительство транспортного узла в какой–либо конкретной точке региона и об оценке срока окупаемости этих инвестиций; показано, что обе задачи, которые формулируются как задачи отыскания максимума функции максимина, сводятся к задаче квадратичного программирования, в то время как отыскание наиболее неблагоприятного сценария формирования рыночных транспортных тарифов на перевозки грузов сводится к решению задач линейного программирования, образующих двойственную пару;

– сформулированы и доказаны проверяемые достаточные условия конкурентоспособности тарифов на услуги транспортного узла в условиях

рынка, которые использованы при формулировке систем ограничений в обеих задачах;

– в рамках исследования теории расписаний в турнирах для случаев трех и четырех участников в одном матче построены обобщения стандартного расписания и расписания равных преимуществ; показано, что данные случаи являются наиболее значимыми с практической точки зрения;

– представлена первая версия концепции платформы для исследования проблем и возможностей по кооперации в Арктике на базе моделей конкурентной кооперации в идеологии многоуровнего подхода к декомпозиции сложных кибер-физических систем;

– в рамках теории многокритериального выбора предложен и исследован подход к анализу чувствительности недоминируемых альтернатив к изменению границ интервалов неопределенности замещения, где альтернативы выбираются на основе интервальных данных оценок замещения критериев;

– в рамках анализа способов выбора оптимальных вариантов и процедур ранжирования альтернатив выполнено эмпирическое исследование устойчивости агрегированных рейтингов, строящихся с помощью ординальных методов, основанных на парных сравнениях по правилу большинства;

– в рамках теории функций доверия исследованы меры конфликта, индуцированные конъюнктивными и дизъюнктивными правилами комбинирования свидетельств; рассмотрены условия согласования правил комбинирования с мерами конфликта;

– в рамках теории функций доверия показано, что в конъюнктивных мерах конфликта можно выделить метрическую и энтропийную составляющие; исследованы свойства изменений этих составляющих при комбинировании свидетельств;

– в рамках исследования степени манипулируемости правил агрегирования предпочтений предложена и исследована модель индивидуального и коалиционного манипулирования правил агрегирования предпочтений в одномерной модели формирования предпочтений; для трех альтернатив выявлены минимально манипулируемые группы мажоритарных правил;

– в рамках анализа в банковской сфере рассмотрено применение модели биномиального теста коррелированных случайных бинарных величин для задачи проверки точности прогноза количества дефолтов;

– проведена оценка влияния мошеннических операций на уязвимость всей финансовой системы; предложен агентно-ориентированный подход моделирования критических уровней мошеннической деятельности, которые делают финансовую систему уязвимой; этот подход применен для исследования таких типов мошенничества, как финансовая пирамида и скрытый риск;

– на основе разработанных ранее моделей влияния в сетях предложен эффективный алгоритм вычисления индексов ближних и дальних взаимодействий; оценена вычислительная сложность этого алгоритма;

– с помощью сетевых моделей проанализированы отношения между странами на различных уровнях взаимодействия, а именно, выполнен сетевой анализ: а) международной миграции с 1970 по 2014гг. на основе данных по потокам миграции между странами; б) торговли продовольствием; в) террористических сообществ;

– проанализирована неоднородность сообщества российских экономистов по результатам анкетирования; предложены несколько методов составления рейтингов научных журналов по экономике с учетом выявленной неоднородности сообщества российских экономистов;

– в рамках исследования паттернов сформулированы и доказаны свойства порядково-инвариантной паттерн-кластеризации; эти свойства показывают, что для практической реализации порядково-инвариантной и диффузионно-инвариантной паттерн-кластеризаций могут быть использованы быстрые алгоритмы сортировки данных;

– в рамках исследования структурной сложности графа предложен новый, вычислительно более мощный вариант частотного алгоритма построения семейства дихотомий произвольного графа, что позволяет вычислить такой показатель сложности графа, как сцепленность; этот алгоритм применен к данным фондового рынка S&P-500 для анализа согласованности его функционирования;

– проведено экспериментальное сравнение алгоритмов кластеризации, основанных на методе аффинного распространения (Affinity Propagation) с другими алгоритмами кластеризации с точки зрения правильности определения количества кластеров и оценки качества полученного разбиения по сравнению с истинным разбиением; численные эксперименты проведены на данных с различной степенью смешивания кластеров; определена иерархия качества алгоритмов кластеризации, основанных на методе аффинного распространения, для разных типов данных;

– разработан вычислительный комплекс, позволяющий моделировать процесс распространения нефти на поверхности Баренцева моря на основе агентно-сетевого подхода;

– исследованы явные двухуровневые конечно-разностные схемы решения одномерных уравнений баротропной газовой динамики; линеаризованы схемы на постоянном решении и выведено необходимое условие типа фон Неймана и критерий типа CFL для слабой консервативности в L^2 для соответствующей задачи на всей прямой;

– предложена и исследована новая фитнес-оптимизационная модель, описывающая влияние факторов окружающей среды на возможности возникновения репродуктивно-соматической специализации в колониях клеток.

Некоторые исследования, выполненные в рамках проекта 2019 года «Исследование моделей принятия решений и анализа сложно структурированных данных», являются продолжением исследований 2018 г. Это, например, исследование свойств методов ранжирования вариантов и правил выбора оптимальных альтернатив, исследование моделей возникновения специализаций в биологических системах, исследование устойчивости сетевых структур, исследование методов анализа связей между различными слоями сетевой структуры, анализ декомпозиционно сложности графов, применение теории функций доверия к задачам социально-экономического анализа и анализа политических систем, применение кластер-анализа и обработки коллекций текстов к задачам структуризации множеств научных публикаций, исследование моделей регулирования Базель II, развитие методологии ранжирования вариантов и правил выбора оптимальных альтернатив с помощью инструментария теории коллективного выбора, исследование новых моделей размещения и др.

Предполагается, что некоторые исследования этого года будут продолжены в дальнейшем.

Методологической базой выполненного проекта были современные методы теории принятия решений, методы оптимизации, теория игр и экономических механизмов, теория графов, теории многокритериального и коллективного выбора, методы анализа данных и машинного обучения, теории вероятностей и неточных вероятностей, теории функций доверия.

Для прикладных исследований предполагается использовать открытые базы данных. Некоторые исследования были выполнены численно с использование готового прикладного программного обеспечения или потребовали разработки комплексов прикладных программ.

Отчет состоит из 6 разделов, соответствующих основным направлениям исследований.

В первом разделе представлены результаты по исследованию моделей выбора и агрегирования. Эти исследования осуществлялись по нескольким направлениям.

В подразделе 1.1 приведены некоторые результаты исследования манипулируемости правил агрегирования предпочтений. Описаны проведенные вычислительные эксперименты для расчета вероятности индивидуального манипулирования при наличии контругроз для 11 правил коллективного выбора. В частности, показано, что правилами подсчета очков с одним скачком баллов всегда можно манипулировать. Кроме того, доказано, что для правила Борда небезопасно манипулируемый профиль существует тогда и только тогда, когда число избирателей или число альтернатив больше или равно 5.

В подразделе 1.2, который носит обзорно-исследовательский характер, рассмотрены различные подходы и методы анализа решений в условиях неопределенности, основанные на теории важности критериев и не предполагающие использования функций полезности и количественных вероятностей.

В подразделе 1.3, который также носит обзорно-исследовательский характер, приведены результаты и выводы по аксиоматическому анализу способов выбора оптимальных вариантов и построения рейтинга альтернатив.

В подразделе 1.4 исследованы группоразделимые свойства профилей предпочтений, а подраздел 1.5 посвящен исследованию стабильных паросочетаний. В частности в нем разработано новое расширение условия последовательного предпочтения, которое приводит к единственному стабильному паросочетанию для любого подмножества агентов.

Второй раздел отчета содержит результаты разработки, исследования и применения методов анализа данных. Исследования проводились в трех областях: в области сетевого анализа данных, анализа текстов и анализа паттернов.

Так, в подразделе 2.1 приведены результаты по сетевому анализу данных. В частности, эти результаты связаны с: разработкой метода оценивания влияния в мультиплексных сетевых структурах, разработкой мер оценивания важности рёбер в сетевых структурах. Кроме того, в подразделе приведены результаты применения разработанных методов к анализу сетей международной торговли продовольствием и торговли вооружением, сети международной миграции.

В подразделе 2.2 в рамках исследования методов анализа текстов вводится понятие оптимального обобщения нечеткого множества тем, сформированного листьями корневого дерева таксономии. На основе этого понятия разработан метод обобщения таксономий, который апробирован на текстовой коллекции научных публикаций по тематике анализа данных, и таксономии Науки данных.

Подраздел 2.3 посвящен описанию новых результатов по развитию методов анализа паттернов. В частности, в этом подразделе описан и исследован алгоритм выделения симметричных паттернов на больших и сильно зашумленных данных.

Третий раздел отчета посвящен описанию результатов в исследования организационных моделей и механизмов. Эти результаты получены в направлениях исследования конкурентной кооперации в Арктике и исследования моделей по оптимальному размещению объектов и привлечению частных инвестиций.

Так, в подразделе 3.1 отчета представлена концепция цикла SATA, которая отражает междисциплинарность применения фундаментальных исследований для улучшения условий жизни в Арктике с использованием передовых технологий.

В подразделе 3.2 описаны результаты, полученные в ходе исследований моделей по оптимальному размещению объектов и привлечению частных инвестиций. В частности, описана и исследована на модельных примерах новая модель для оценки необходимого объема инвестиций в развитие региональной

инфраструктуры грузоперевозок; выявлены и описаны классы системных проблем, связанных с формированием государственно-частного партнерства (на примере энергетических систем); модифицированы математические модели, на основе которых переформулированы математические задачи оценки объема инвестиций, требуемого для размещения цепи зарядных станций электромобилей на шоссе.

В четвертом разделе отчета приведены результаты исследований моделей анализа политических систем. Эти исследования проходили в направлениях анализа: а) значимости позиций партий для голосования и б) поляризованности и влияния в политических системах.

Так подраздел 4.1 отчета посвящен исследованию одной модели анализа значимости позиций партий для голосования по ним методами теории монотонных мер и теории функций доверия.

В подразделе 4.2 проанализированы основные модели и методы оценивания поляризации и политического влияния, в том числе и модели, учитывающие предпочтения по формированию коалиций.

Пятый раздел посвящен описанию исследований моделей и анализа данных в финансовой и банковской сферах.

Так в подразделе 5.1 приведены результаты исследований влияния модельного риска на точность оценок величины риск-взвешенных активов, а в подразделе 5.2 приведены результаты анализа структурной сложности данных по фондовому рынку S&P-500.

Наконец шестой раздел посвящен исследованию моделей распространения и специализации.

Так, в подразделе 6.1 приведены результаты исследований явных двухуровневых конечно-разностных схем решения полной системы уравнений газовой динамики для случая произвольной фоновой скорости, а также результаты разработки универсальной сетевой модели нахождения оптимального расположения сенсоров в море для обнаружения загрязнений.

В подразделе 6.2 приведены некоторые новые результаты исследования оптимизационной задачи возникновения специализации в абстрактной многоклеточной системе.

Приведенные в отчете результаты были апробированы на семинарах, представлены на конференциях различного уровня, в том числе и на международных научных конференциях. Большинство описанных в отчете результатов опубликованы, поданы или подготовлены к публикации.

Выполнение проекта «Исследование моделей принятия решений и анализа сложно структурированных данных» осуществлялось в 2019 г. в следующих научных направлениях:

– исследование моделей выбора и агрегирования;

– разработка, исследование и применение методов анализа данных;

– исследование организационных моделей и механизмов;

– исследование моделей анализа политических систем;

– исследование моделей в финансовой и банковской сферах;

– исследование моделей распространения и и специализации.

1 Исследование моделей выбора и агрегирования

В этом разделе отчета приведены результаты исследований моделей выбора и агрегирования. Работа проводилась в следующих направлениях:

– построение и применение теоретических моделей для расчета степени манипулируемости правил агрегирования предпочтений;

– разработка методов анализа решений при нечисловом оценивании предпочтений и неопределенности на основе теории важности критериев;

– аксиоматический анализ способов выбора оптимальных вариантов;

– исследование свойств профилей предпочтений.

1.1 Построение и применение теоретических моделей для расчета степени манипулируемости правил агрегирования предпочтений

1.1.1 Введение

Проблема принятия коллективных решений широко известна: она возникает во всех ситуациях, когда требуется учесть мнение множества агентов, но решение необходимо принять одно. Результат, т.е. коллективное решение, можно получить, собрав предпочтения всех избирателей и агрегировав их при помощи некоторого правила коллективного выбора. Существует множество процедур, отличных по своим свойствам, однако для любого однозначного недиктаторского правила доказано, что, если имеется хотя бы три возможных исхода, то правило подвержено манипулированию (теорема Гиббарда-Саттертуэйта [2, 3]). Это означает, что избиратели могут добиться более предпочтительного для них результата голосования, намеренно исказив свои предпочтения.

В рамках проекта МЛАВР в 2019 году исследование манипулирования было продолжено в двух направлениях: экспериментальное изучение индивидуального манипулирования при контругрозе и теоретическое исследование безопасного коалиционного манипулирования.

1.1.2 Манипулирование при контругрозе

1.1.2.1 Постановка задачи

Была поставлена задача изучить проблему манипулирования в ситуации, когда не только один участник голосования пытается сообщить во время голосования неискренние предпочтения, но и когда какой-то другой участник может ответить на данную попытку манипулирования, также сообщив неискренние предпочтения. Данная концепция названа манипулирование при контругрозе.

Манипулирование при контругрозе рассматривалось в литературе в нескольких источниках, один из наиболее полных – статья [4]. В этой статье рассматривалось сразу несколько разных случаев: и простое манипулирование, и ответы других участников в виде контругроз.

Мы рассматриваем 11 правил коллективного выбора для 3 и 4 альтернатив. Мы рассматриваем алфавитный принцип устранения несравнимости при голосовании, а также концепцию множественного выбора и расширенных предпочтений. Мы рассматриваем проблему одиночного манипулирования и контругрозы.

1.1.2.2 Правила коллективного выбора

Были рассмотрены 11 правил коллективного выбора:

1) правило относительного большинства (Plurality Rule, Pl);
2) одобряющее голосование, когда каждый участник выбирает ровно двух кандидатов (2-Approval rule, A2);
3) правило Борда (Borda’s Rule, Bo);

4) процедура Блэка (Black’s Procedure, Bl);

5) правило Коупланда I (C1);

6) правило Коупланда II (C2);

7) правило Коупланда III (C3);

8) пороговое правило (Threshold rule, T);

9) процедура Нансона (Nanson’s Procedure, N);

10) обратная процедура Борда (Inverse Borda’s Procedure, InvBo);

11) процедура Хара (Hare’s Procedure, H).

В литературе можно встретить немало работ, в которых оценивается степень одиночного манипулирования этих правил коллективного выбора, например, [5, 6]. Мы используем определения правил, изложенные в данных источниках.

Были рассмотрены случаи Impartial Culture и Impartial Anonymous Culture. Рассматривалось 3 и 4 кандидата в голосовании, а также от 3 до 30 участников голосования.

1.1.2.3 Расширенные предпочтения, множественный и одиночный выбор

Были рассмотрены несколько моделей для представления итогов голосования. В случае, если по итогам правила коллективного выбора две альтернативы показывали одинаковое количество голосов или набирали одинаковое количество баллов, мы использовали как концепцию алфавитного выбора, так и концепцию множественного выбора.

В случае с алфавитным выбором при образовании ничьи между несколькими альтернативами выбирается та, которая идёт раньше по алфавиту.

В случае с множественным выбором полагается, что в результате получается ничья (множественный выбор), а у участника голосования есть расши-

ренные предпочтения на множестве всех возможных исходов голосования, которые позволяют ему понять, какой исход из нескольких будет более привлекательным.

Например, для трёх альтернатив использовались 4 различных метода построения расширенных предпочтений. Предполагая, что для участника голосования альтернативы упорядочены как $a \succ b \succ c$, получаем, что в случае Лексимина (Leximin) расширенные предпочтения будут выглядеть как: $\{a\} \succ \{a,b\} \succ \underline{\{b\} \succ \{a,c\} \succ \{a,b,c\}} \succ \{b,c\} \succ \{c\}$. Для Лексимакса (Leximax): $\{a\} \succ \{a,b\} \succ \underline{\{a,b,c\} \succ \{a,c\} \succ \{b\}} \succ \{b,c\} \succ \{c\}$. Для расширения «По возрастанию вероятности наихудшего» (PWorst): $\{a\} \succ \{a,b\} \succ \underline{\{b\} \succ \{a,b,c\} \succ \{a,c\}} \succ \{b,c\} \succ \{c\}$. По «Убыванию вероятности наилучшего» (PBest): $\{a\} \succ \{a,b\} \succ \underline{\{a,c\} \succ \{a,b,c\} \succ \{b\}} \succ \{b,c\} \succ \{c\}$.

1.1.2.4 Схема расчетов

Для оценки степени манипулируемости был использован индекс Нитцана-Келли (NK), который используется во многих работах по манипулированию, например, в [5, 6]. Индекс Нитцана-Келли – это доля манипулируемых профилей (ситуаций) к общему количеству профилей.

Мы использовали компьютерное моделирование для оценки значений индекса NK для разных случаев. Мы генерировали не менее 10 тысяч профилей для каждого случая, после чего рассчитывали, существует ли хотя бы одна ситуация, когда он был манипулируем.

Схема была следующей:

1) генерируется не менее 10 тысяч профилей;

2) для каждого профиля высчитывается коллективный выбор по каждому правилу;

3) для каждого профиля проверяется, есть ли хотя бы один участник, который может хотя бы одним способом изъявить неискренние предпочтения, чтобы новый результат оказался для него лучше первоначального;

4) если существует хотя бы один такой профиль, то этот профиль является манипулируемым с точки зрения индивидуального манипулирования;

5) в случае с контругрозами дополнительно проверяется, нет ли возможности у кого-то другого в этом профиле ответить на попытку манипулирования. И если для каждой успешной попытки манипулирования существует не менее одной ситуации, когда какой-то другой участник голосования также изъявляет неискренние предпочтения, чтобы сделать выбор для себя не хуже первоначального, а для манипулирующего участника – не лучше первоначального, то в этом случае считается, что за счет контругрозы профиль неманипулируем.

Очевидно, что манипулируемость при контругрозе ниже, чем при обычном индивидуальном манипулировании, когда ответные ходы (контругрозы) не предусмотрены.

1.1.2.5 Результаты

В качестве базовой статьи для терминологии по теме контругроз была взята работа [4]. Рассчитываемый нами сценарий совпадает с 4-м типом манипулирования, который изложен в первоначальной статье.

В отличие от первоисточника, мы рассматриваем случай с контругрозами не только для алфавитного случая несравнимости, но и для расширенных предпочтений Leximin, Leximax, PWorst и PBest.

Часть результатов, полученных нами, можно увидеть на рисунках 1.1 и 1.2.

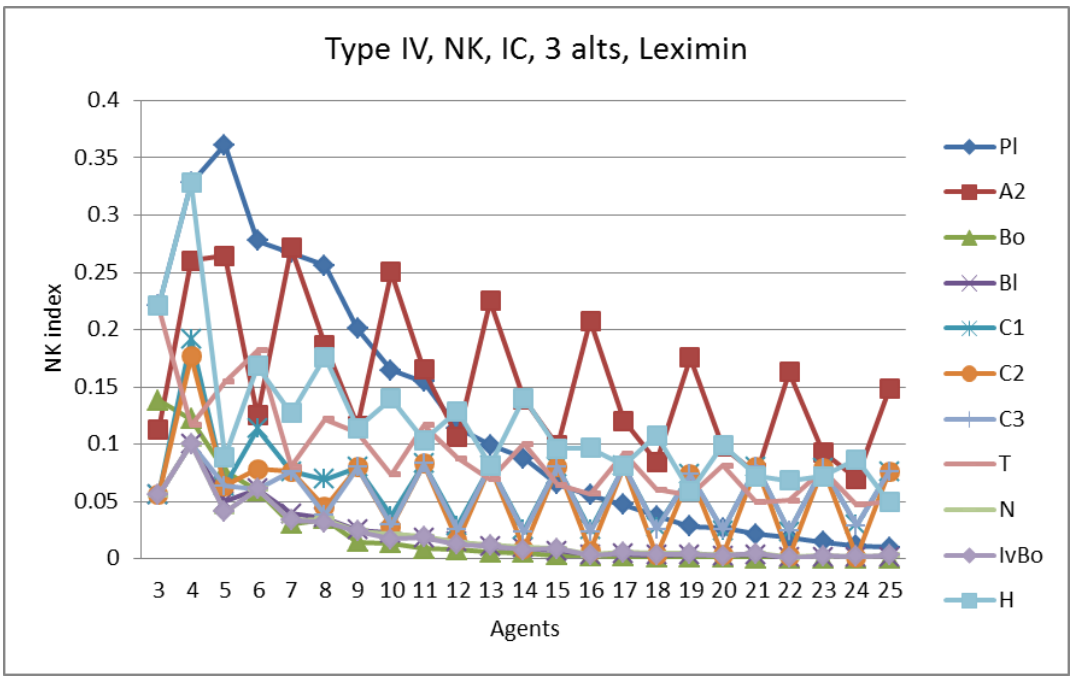


Рисунок 1.1 – Манипулирование для случая с контругрозами, 3 альтернативы, Impartial Culture, Leximin

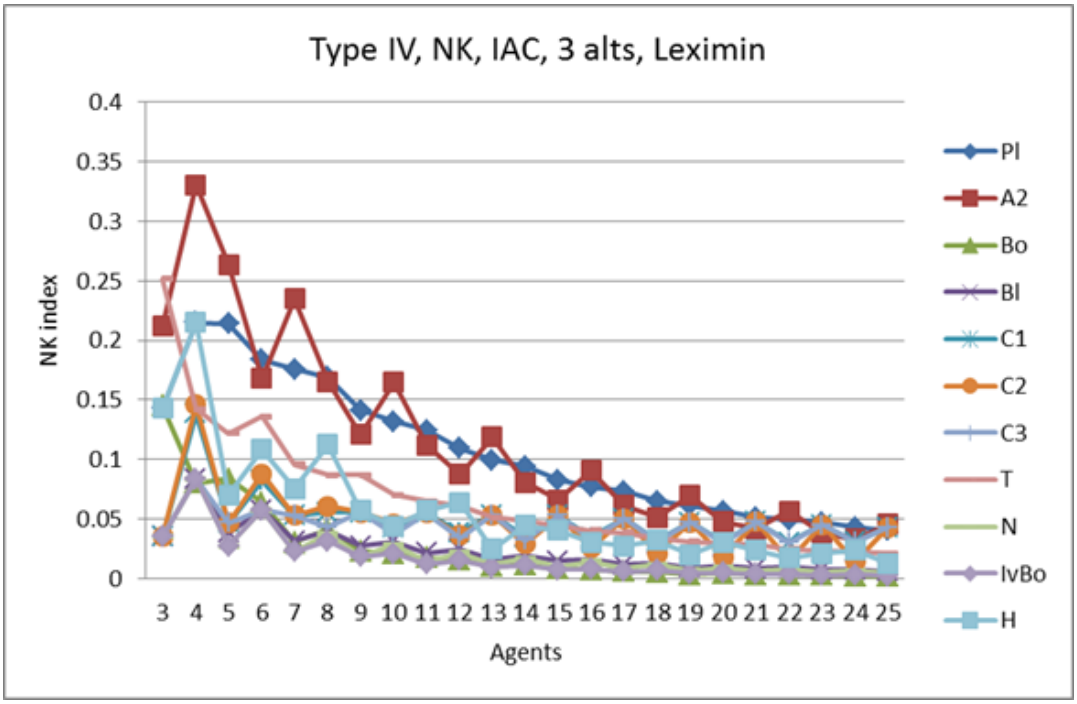


Рисунок 1.2 – Манипулирование для случая с контругрозами, 3 альтернативы, Impartial Anonymous Culture, Leximin

Сразу видны несколько правил, которые для обоих случаев являются наименее манипулируемыми. Рассмотрим их отдельно в чуть большем масштабе (рисунки 1.3 и 1.4).

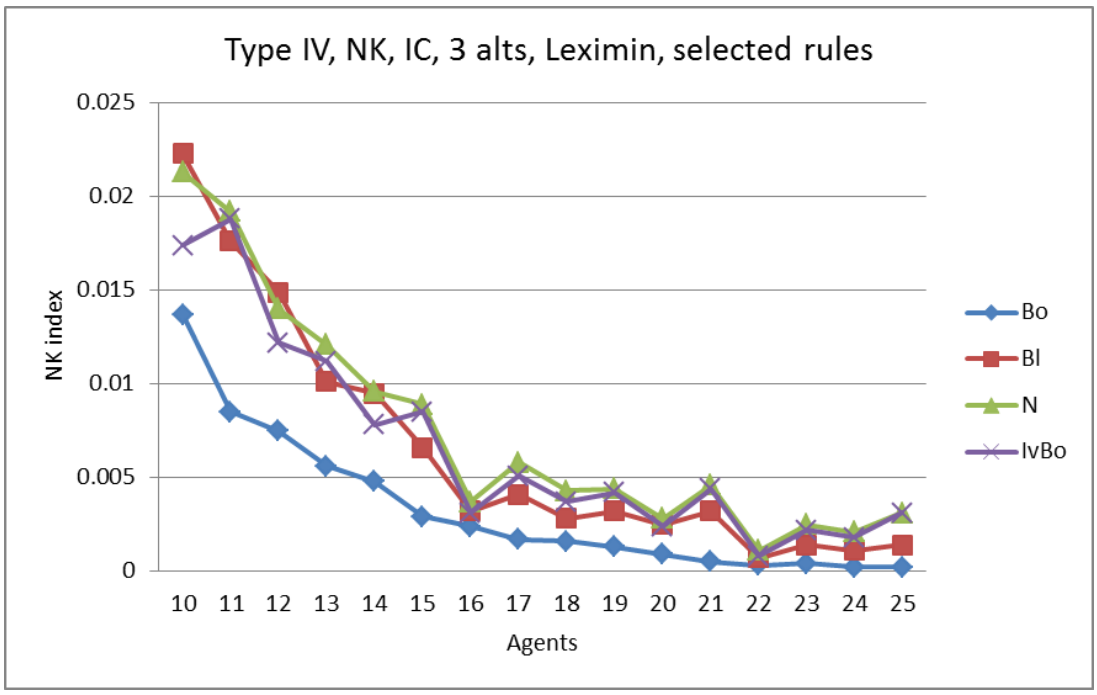


Рисунок 1.3 – Наименее манипулируемые правила, Impartial Culture, Leximin

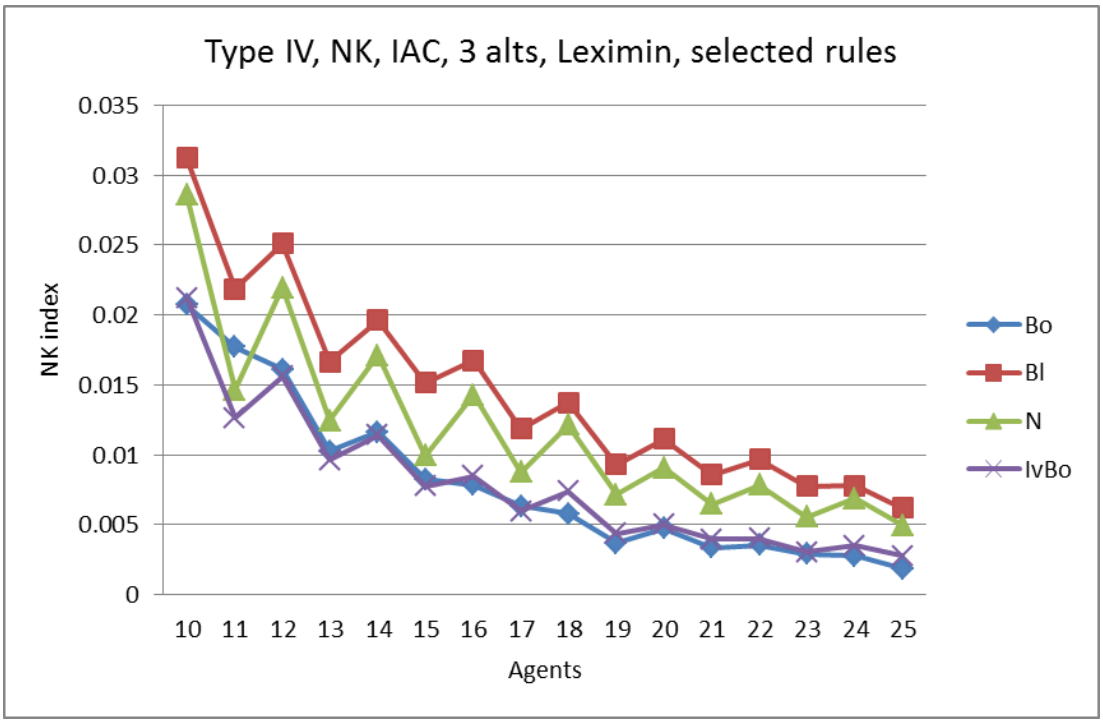


Рисунок 1.4 – Наименее манипулируемые правила, Impartial Anonymous Culture, Leximin

Процедура Борда (Borda’s Rule) оказалась наименее манипулируемой для большинства рассмотренных случаев, показывая наименьшие значения индекса NK. Обратная процедура Борда (Inverse Borda’s Procedure) и правило Нансона (Nanson's Procedure) показали близкие, но чуть худшие результаты

для большинства случаев. При этом схожие выводы для алфавитного устранения несравнимости можно найти в [4], а мы их получили и для нескольких разных случаев множественного выбора.

При этом правило Борда не является наименее манипулируемым для классического случая манипулирования, когда нет контругроз. Объяснением результата может быть гибкость, которой обладает участник при голосовании в процедуре Борда. Так как каждый участник может располагать в бюллетене кандидатов разными способами, и они будут получать за это разные баллы, то способов оказать влияние на исход голосования в среднем больше, чем по другим правилам. Однако и контругроз в таком случае, то есть ответов других участников будет больше. В результате правило Борда не является самым хорошим для случая индивидуального манипулирования [6], однако является таковым для случая с контругрозами.

1.1.3 Безопасное коалиционное манипулирование

1.1.3.1 Постановка задачи

В рамках проекта рассматривалась также проблема манипулирования результатом голосования со стороны коалиций избирателей. Коалицией называется группа избирателей, имеющих одинаковые предпочтения. Если в результате изменения предпочтений на неискренние всеми членами коалиции удается достичь более предпочтительного для них результата голосования, то каждый из членов коалиции имеет стимул манипулировать. В то же время, при недостаточной согласованности действий членов коалиции существует опасность получить результат худший, чем был изначально, если не все члены коалиции решают манипулировать. В этом случае коалиционное манипулирование небезопасно.

Проблема несогласованности стратегических действий групп избирателей исследовалась также в [7], однако предложенное авторами определение безопасного манипулирования несколько отличается от рассматриваемого в

данной работе. Предполагается, что избиратель имеет стимул манипулировать, если существует некоторое подмножество избирателей из множества его единомышленников (коалиции), которые при одинаковом изменении предпочтений могут добиться более выгодного для себя результата голосования. При такой формулировке существует проблема координирования действий членами коалиции: кто из них должен манипулировать, и каким именно образом, а кто – нет. Поэтому авторами [7] предложено считать манипулирование небезопасным, если существует возможность получить результат хуже первоначального, если слишком мало или слишком много членов коалиции решат манипулировать.

Предложенное в данной работе определение манипулирования в большей степени подходит для ситуаций, когда коммуникационные возможности членов коалиции ограничены. Каждый избиратель может предполагать, что обладающие идентичными предпочтениями другие избиратели будут действовать идентично, поэтому принимает решение о манипулировании, учитывая, что его единомышленники также будут манипулировать. С другой стороны, избирателю необходимо понять, не приведет ли попытка манипулирования к провалу в случае, если некоторые члены коалиции откажутся манипулировать, несмотря на то, что это более выгодно. Введение этих предпосылок является следующим шагом в моделировании независимо действующих избирателей после наиболее простой модели индивидуального манипулирования, в которой избиратели не учитывают стимулы других избирателей к манипулированию.

В рамках проекта получены результаты теоретического исследования небезопасного коалиционного манипулирования в профиле, если применяется правило подсчета очков. Это исследование вносит вклад в литературу по данной теме, так как ранее рассматривался только вопрос о существовании безопасного манипулирования. Однако безопасно манипулируемый профиль существует во всех случаях, когда правило удовлетворяет условиям теоремы

Гиббарда-Саттертуэйта [7], а небезопасное манипулирование имеет место далеко не всегда.

1.1.3.2 Примеры, термины и обозначения

1.1.3.2.1 Пример небезопасного манипулирования

Рассмотрим пример небезопасного коалиционного манипулирования. Профиль предпочтений для 7 избирателей представлен в следующей Таблице 1.1. Первый столбец Таблицы 1.1 означает, что имеются три избирателя, для которых альтернатива *a* является наилучшей, *b* – второй наилучшей и т.д.

Таблица 1.1 – Профили предпочтений для 7 избирателей.

3	**2**	**1**	**1**
a	d	d	e
b	c	c	d
c	b	e	a
e	e	a	c
d	a	b	b

Пусть для принятия коллективного решения используется правило Борда. Согласно этому правилу, за каждое k-ое место в индивидуальных предпочтениях альтернатива получает m – k очков. Сумма очков по всем избирателям для альтернативы x обозначается за S(x). Таким образом, суммарное количество очков для каждой альтернативы: S(a)=15, S(b)=13, S(c)=16, S(d)=15, S(e)=11. Победитель голосования – альтернатива c, набравшая наибольшее количество очков. Рассмотрим коалицию K, состоящую из трех избирателей первого столбца таблицы, имеющих предпочтения aPbPcPePd. Члены коалиции K имеют стимул манипулировать, сообщая неискренние предпочтения bPaPePcPd. В этом случае распределение очков следующее: S(a)=12, S(b)=16, S(c)=13, S(d)=15, S(e)=14. Результат голосования – альтернатива b, являющаяся более предпочтительной для членов коалиции. Теперь предположим, что манипулировать решает только один избиратель из коалиции K. Тогда S(a)=14, S(b)=14, S(c)=15, S(d)=15, S(e)=12. Наибольшее количество очков набирают альтернативы c и d. Кроме того, предположим, что по правилу устранения множественности выбора выбирается альтернатива d. Таким образом, результат голосования оказывается для членов коалиции менее предпочтительным,

чем был изначально. Манипулирование со стороны членов коалиции с использованием стратегии bPaPePcPd называется небезопасным.

1.1.3.2.2 Правила подсчета очков

Обозначим за m число альтернатив, за n – количество избирателей. Ниже будут представлены результаты исследования безопасности коалиционного манипулирования для семейства правил подсчета очков, которому принадлежит правило относительного большинства и уже упомянутое правило Борда. Любое правило из этого семейства определяется вектором очков $s = (s_1 \geq \cdots \geq s_m)$, $s_1 \geq \cdots \geq s_m$ где s_j обозначает количество очков, присваиваемых альтернативе за j-ю позицию в индивидуальных предпочтениях избирателя. Общее количество очков для альтернативы a обозначается за $S(a)$ и вычисляется как сумма очков по всем избирателям. Примеры векторов очков:

– правило относительного большинства: s_{Pl} = (1, 0, ..., 0);

– правило вето: s_V = (1, ..., 1, 0);

– правило Борда: s_B = (m − 1,m − 2, ..., 1, 0).

Ясно, что для любого правила подсчета очков хотя бы одно нестрогое неравенство должно выполняться как строгое неравенство. Пусть

$k_1 = min(j \in \{1, \dots, m\}, \text{т.ч } s_j - s_{j+1} > 0)$, $\Delta_1 = s_{k_1} - s_{k_1+1}$,

$k_2 = min\big(j \in \{k_1, \dots, m\}, \text{т.ч } s_j - s_{j+1} > 0\big)$, $\Delta_2 = s_{k_2} - s_{k_2+1}$,

$k_3 = min\big(j \in \{k_2, \dots, m\}, \text{т.ч } s_{j+1} - s_j > 0\big)$, $\Delta_3 = s_{k_2} - s_{k_2+1}$, и т.д.

Таким образом, k_l – номер элемента вектора очков, для которого следующий за ним элемент меньше данного (имеет место скачок значений элементов вектора). Пусть q – количество таких элементов (скачков).

1.1.3.3 Результаты

Теорема 1.1. Условия существования небезопасно манипулируемого профиля для правил подсчета очков при $n \geq 3$ приведены в Таблице 1.2.

Таблица 1.2 – Условия существования небезопасно манипулируемого профиля для правил подсчета очков при $n \geq 3$

	q = 1	q = 2		q ≥ 3
		$\Delta_1 > \Delta_2$	$\Delta_1 \leq \Delta_2$	
m = 3	Нет	Нет	Нет	-
m = 4	Нет	Да	Нет	Да, если $\Delta_1 > \Delta_3$ или $\Delta_2 > \Delta_3$, Нет, иначе.
m = 5	Нет	Да	Нет, если $k_1 = 2$ и $k_2 = 4$, Да, иначе	Да
$m \geq 6$	Нет	Да	Да	Да

Как следствие из теоремы 1.1, любое коалиционное манипулирование для правила относительного большинства и правила вето является безопасным для любых $n \geq 3$ и $m \geq 3$, так как в векторе очков этих правил только один скачок значений. Для правила Борда небезопасно манипулируемый профиль существует при $m \geq 5$, так как после все значения элементов вектора очков правила Борда различны, и $\Delta_1 = \Delta_2 = \Delta_3 \ldots$. Точное значение минимального числа избирателей, необходимого для существования небезопасного манипулирования определяется следующей теоремой.

Теорема 1.2. Небезопасно манипулируемый профиль для правила Борда существует, если и только если $\mathrm{m} \geq 5$ и $\mathrm{n} \geq 5$.

Таким образом, определены условия, при которых существует небезопасное коалиционное манипулирование для правил подсчета очков.

1.1.4 Некоторые выводы по исследованию манипулирования

В исследовании манипулирования со стороны избирателей в 2019 году были сделаны важные шаги в приближении моделей к реальным ситуациям принятия коллективных решений. Изучение индивидуального манипулирования при контругрозах и коалиционного безопасного манипулирования является попыткой понять, как меняются стимулы избирателей к манипулированию, если они учитывают возможность других избирателей выбирать стратегию голосования.

Введение предпосылки о контругрозах в случае некоторых правил имеет решающее значение. Так, процедура Борда показывает высокие показатели

манипулируемости при простом индивидуальном манипулировании, однако она же оказалась наименее манипулируемой для большинства рассмотренных случаев при контругрозах.

Аналогично, если коалиционное манипулирование не является безопасным, то это создает отрицательный стимул к искажению предпочтений. Возможно, доля таких профилей столь невелика, что небезопасность манипулирования не является фактическим препятствием для коалиционного манипулирования. Однако данный вопрос пока остается открытым и является возможной темой дальнейших исследований в этом направлении.

Результаты исследований отражены в работе [8, 9, 10, 11].

1.2 Разработка методов анализа решений при нечисловом оценивании предпочтений и неопределенности на основе теории важности критериев

1.2.1 Введение

Практические задачи принятия решений часто предполагают необходимость учета неопределенных факторов, значения которых заранее точно неизвестны. Основные, наиболее распространенные подходы к анализу таких задач предполагают использование функции полезности, т.е. функции, математическое ожидание которой моделирует предпочтения ЛПР (лица, принимающего решение) на множестве распределений вероятностей на множестве исходов [12, 13]. Однако построение функции полезности – проблема весьма сложная и потому не всегда решаемая [14]. И даже само допущение существования функции полезности является достаточно сильным, так как может привести к появлению разного рода «парадоксов» – выводов, не согласующихся со здравым смыслом или очевидно ожидаемых ЛПР результатов [15,16]. С другой стороны, использование функции полезности предполагает, что задано (полностью или хотя бы частично) распределение вероятностей на множестве зна-

чений неопределенного фактора. Однако, в сложных задачах принятия решений социально-экономического, политического и иного характера объективных вероятностей обычно не существует и приходится обращаться к экспертным оценкам. Но, как известно, количественные оценки весьма сложны для человека, а потому являются недостаточно надежными и могут содержать явные ошибки [16].

Поэтому теоретический и практический интерес представляет проблема разработки методов анализа решений в условиях неопределенности, которые не использовали бы функций полезности и количественных вероятностей. Оказывается, что ряд таких методов можно построить на основе результатов теории качественной важности критериев [17, 18], разработанной для анализа многокритериальных задач, так как понятие важности критериев оказывается в некоторых отношениях формально схоже с понятием вероятности. На это обстоятельство ранее уже указывалось в [19, 20].

Первые результаты приложения теории важности критериев к задачам принятия решений при неопределенности были получены в [21]. В данной части отчета, являющейся, по сути, обзорно-исследовательской, подходящие определения и методы теории качественной важности изложены сразу в терминах анализа решений при неопределенности. Кроме того, представлены и новые результаты, которые, в свою очередь, приложимы и к многокритериальным задачам. Теоретические положения иллюстрируются простыми расчетными примерами.

1.2.2 Математическая модель проблемной ситуации

Далее используется следующая математическая модель ситуации принятия решения в условиях неопределенности:

$$< \eta, S, \Lambda, \Theta, f, \Pi, R, \Gamma >. \qquad (1.1)$$

Здесь:

η – тип постановки задачи (выбрать одну наилучшую стратегию (η_1) или же l лучших (η_l), упорядочить все стратегии по предпочтительности ($\eta_{\rightarrow}$) и т.д.);

S – множество стратегий (планов, альтернатив, вариантов, действий, решений, …);

$\Lambda = \{\lambda^1, \lambda^2, \ldots, \lambda^m\}$ – конечное множество значений неопределенного фактора ($m \geqslant 2$); его элементы называются «состояниями природы», или элементарными событиями, а его подмножества именуются событиями;

Θ – числовая модель оценивания предпочтительности последствий. Она включает в свой состав множество V числовых оценок последствий: $V \subseteq (-\infty, +\infty)$. Предполагается, что предпочтения возрастают на V, т.е. чем оценка v больше, тем она предпочтительнее. Множество V может быть континуальным (например, $V = [v_*, v^*]$ для денежных выигрышей) или дискретным (например, для пятибалльной целочисленной шкалы $V = \{1, 2, 3, 4, 5\}$). Если специально не указано иное, то считается, что предпочтения оцениваются в порядковой шкале[1] [22]. Но шкала может быть и более совершенной.

f – функция выигрыша (критерий, показатель эффективности или качества, и т.п.) – отображение $S \times \Lambda \rightarrow V$, т.е. $v = f(s, \lambda)$ – оценка последствия использования стратегии s, если реализуется значение неопределенного фактора λ. То есть V – это область значений функции f. Каждая стратегия s характеризуется вектором (векторной оценкой последствий) $x = (f(s, \lambda^1), f(s, \lambda^2), \ldots, f(s, \lambda^m))$. Множество всех векторных оценок x (как достижимых, т.е. соответствующих стратегиям и значениям неопределенного фактора, так и гипотетических) есть $X = V^m$.

---

[1] Шкала предполагается порядковой с точки зрения моделирования предпочтений. Однако «физическая» шкала может быть количественной. Так, для денег имеем «физическую» шкалу отношений, так что, например, можно сказать, что сумма в 20 рублей в 2 раза больше, чем 10 рублей, но при этом нельзя утверждать, что она предпочтительнее в 2 раза.

П – нечисловая модель вероятностей. Она содержит (бинарное) отношение (нестрогого превосходства в) вероятности $\propto$ на множестве $L = 2^{\Lambda}$ подмножеств множества Λ: запись $A \propto B$ означает, что событие A не менее вероятно, чем B. Кроме этого, модель П может содержать и иные отношения.

R – отношение нестрогого предпочтения ЛПР на множестве X: запись xRy означает, что вектор x не менее предпочтителен, чем y. Оно порождает отношение (строгого) предпочтения P и безразличия (равенства по предпочтительности) I: xPy верно, когда верно xRy и неверно yRx, и xIy верно, когда верно и xRy, и yRx. Запись xPy означает, что вектор x предпочтительнее, чем y, а xIy означает, что векторы x и y безразличны (одинаковы по предпочтительности). Иными словами, отношения I и P являются симметричной и асимметричной частями отношения R, т.е. $I = \mathrm{Sym}\, R = R \cap R^{-1}$, $P = \mathrm{As}\, R = R \setminus \mathrm{Sym}\, R$, где R^{-1} – отношение, обратное к R. Предполагается, что отношение R есть квазипорядок (оно рефлексивно: xRx верно для любого $x \in X$, и транзитивно: для любых $x, y, z \in X$ из xRy и yRz следует xRz), но лишь частичный, т.е. не всякие векторы x и y из X сравнимы по R: может не выполняться ни xRy, ни yRx. Для квазипорядка R отношение P оказывается строгим частичным порядком (оно иррефлексивно: yPy неверно для любого $y \in Z$, и транзитивно), а отношение I – эквивалентностью (оно симметрично: для любых $x, y \in X$ из xRy следует yRx, рефлексивно и транзитивно). Отношения R, P, I порождают на множестве стратегий S аналогичные по смыслу отношения R_*, P_*, I_* следующим образом:

$$s'R_*s'' \Leftrightarrow x'Rx'',\ s'P_*s'' \Leftrightarrow x'Px'',\ s'I_*s'' \Leftrightarrow x'Ix'',$$

где $x' = (f(s', \lambda^1), f(s', \lambda^2), \ldots, f(s', \lambda^m))$ и $x'' = (f(s'', \lambda^1), f(s'', \lambda^2), \ldots, f(s'', \lambda^m))$. Именно отношения R_*, P_*, I_* непосредственно используются для формирования решения анализируемой задачи принятия решения с учетом

типа ее постановки. Например, если требуется выбрать одну наилучшую (оптимальную) стратегию, то строится множество недоминируемых по P_* стратегий, среди которых (если множество P_* внешне устойчиво) и надлежит выделить наилучшую.

Г – информация об отношении ЛПР к риску. Оно может быть не склонным к риску ($\Gamma = \{\gamma^1\}$), склонным к риску ($\Gamma = \{\gamma^2\}$), или безразличным (нейтральным) к риску ($\Gamma = \{\gamma^3\}$). Но такой информации может и не быть ($\Gamma = \varnothing$).

Отношение R заранее неизвестно: его нужно построить на основе модели оценивания предпочтительности последствий Θ, модели вероятностей неопределенного фактора П и информации об отношении ЛПР к риску Г.

1.2.3 Отношения вероятности

Отношение $\propto$ порождает отношения превосходства по вероятности ϕ и равновероятности ~: $A\phi B$ верно, когда верно $A \propto B$ и неверно $B \times A$, и $A \sim B$ верно, когда верно и $A \propto B$, и $B \times A$. Запись $A\phi B$ означает, что событие A более вероятно, чем B, а запись $A \sim B$ означает, что события A и B равновероятны.

Выделим следующие важные частные виды отношения вероятности ϕ, которые будут использоваться в дальнейшем.

Полная качественная вероятность. Отношение $\propto$ называется (*полной*) *качественной* (или *сравнительной*) *вероятностью*, если оно обладает следующими свойствами (удовлетворяет следующим аксиомам де Финетти (1931)) [23, 24]:

A1. Сравнимость: для любых A и B выполнено, по крайней мере, одно из соотношений: $A \propto B$ или $B \times A$.

A2. Транзитивность: из соотношений $A \propto B$ и $B \propto C$ следует $A \propto C$.

A3. Существенность: $\Lambda \propto \varnothing$ и для любого A выполнено $A \propto \varnothing$.

А4. Аддитивность: если $A \cap C = B \cap C = \varnothing$, то соотношения $A \propto B$ и $A \cup C \propto B \cup C$ выполнены или не выполнены одновременно.

Отметим, что из А1 и А2 следует рефлексивность.

Пусть на Λ задана количественная вероятность Pr рядом распределения

$$\Pr(\lambda^1) = p_1, \ldots, \Pr(\lambda^m) = p_m,$$

так что вероятность события A равна:

$$\Pr A = \sum_{j:\, \lambda^j \in A} p_j .$$

Говорят, что количественная вероятность численно представляет качественную вероятность, или, что количественная вероятность согласована с качественной, если выполнено условие: для любых событий A и B

$$A \propto B \Leftrightarrow \Pr A \geqslant \Pr B. \tag{1.2}$$

Если задана количественная вероятность Pr, то порождаемое ею согласно (1.2) отношение $\propto$ удовлетворяет всем аксиомам А1 – А4, т.е. является полной качественной вероятностью. Оказывается, однако, что предположение о том, что всякую полную качественную вероятность можно представить некоторой количественной, неверно! Контрпример был построен в [25]. В свете этого утверждения становится понятным, что методы, опирающиеся на предположение о существовании количественной вероятности (в том числе методы теории полезности), при наличии качественной вероятности в принципе не применимы.

Частичная качественная вероятность. Отношение $\propto$ называется частичной качественной (или частичной сравнительной) вероятностью [26, 27], если оно удовлетворяет аксиомам А2 – А4 и аксиоме

А5. Рефлексивность: для любого A выполнено $A \propto A$.

Порядковая вероятность. Отношение $\propto$ называется частичной порядковой (или частичной ординальной) вероятностью, если оно упорядочивает по

вероятности лишь отдельные элементарные события и является частичным квазипорядком. Если такое отношение упорядочивает по вероятности все элементарные события, то оно называется (полной) порядковой (или ординальной) вероятностью. Если перенумеровать элементарные события в порядке невозрастания их вероятности, то (полную) порядковую вероятность можно кратко представить так:

$$\{\lambda^1\} \sim \ldots \sim \{\lambda^{i_1}\} \succ \{\lambda^{i_1+1}\} \sim \ldots \sim \{\lambda^{i_2}\} \succ \ldots \succ \{\lambda^{i_{\tau-1}+1}\} \sim \ldots \sim \{\lambda^{i_\tau}\} \qquad (1.3)$$

(очевидно, что $i_\tau = m$). Таким образом, элементарные события объединяются в τ групп так, что элементарные события из одной и той же группы равновероятны, а элементарные события из группы с меньшими номерами вероятнее элементарных событий из групп с большими номерами. Пусть M_t – множество номеров элементарных событий из t-й группы, $t = 1, 2, \ldots, \tau$.

1.2.4 Построение отношений предпочтения и безразличия

Разберем, как строится отношение R на основе информации о неопределенном факторе – с использованием вероятности $\propto$. При наличии информации об отношении ЛПР к риску ее также следует использовать.

Для векторов и матриц одинаковых размерностей введем обозначения:

$a \geqq b \Leftrightarrow a_i \geqslant b_i$, $i = 1, 2, \ldots, n$; $a \geq b \Leftrightarrow (a \geqq b, a \neq b)$; $a > b \Leftrightarrow a_i > b_i$, $i = 1, 2, \ldots, n$ (здесь $n > 1$).

$$\|a_{ij}\| \geqq \|b_{ij}\| \Leftrightarrow a_{ij} \geqslant b_{ij},\ i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, n;$$

$$\|a_{ij}\| \geq \|b_{ij}\| \Leftrightarrow (\|a_{ij}\| \geqq \|b_{ij}\|,\ \|a_{ij}\| \neq \|b_{ij}\|),$$

$$\|a_{ij}\| > \|b_{ij}\| \Leftrightarrow a_{ij} > b_{ij},\ i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, n \text{ (здесь } m > 1 \text{ или } n > 1).$$

Поскольку с увеличение оценки v в V предпочтения возрастают, то на множестве X оказывается определенным отношение Парето $R^{\varnothing}$: $xR^{\varnothing}y \Leftrightarrow x \geqq y$. Заметим, что $xP^{\varnothing}y \Leftrightarrow x \geq y$, а $I^{\varnothing}$ есть отношение равенства векторов =.

Пусть A и B – два сравнимых по вероятности $\propto$ события, т.е. верно $A{\sim}B$, $A{\propto}B$ или $B{\times}A$. Без ограничения общности примем, что они несовместны, т.е. $A \cap B = \varnothing$ (в противном случае будем рассматривать события $A\backslash C$ и $B\backslash C$, где $C = A \cap B$). Кроме того, будем полагать, что множества A и B не пусты. Рассмотрим векторную оценку $x^{AB} = (x_1, \ldots, x_m)$, в которой все компоненты с номерами i, для которых $\lambda^i{\in}A$, равны между собой, и все компоненты с номерами j, для которых $\lambda^j{\in}B$, также равны между собой. Обозначим через $x^{A\leftrightarrow B}$ вектор, полученный из x^{AB} заменой каждой компоненты с номером i, для которого $\lambda^i{\in}A$, (любой) компонентой с номером j, для которого $\lambda^j{\in}B$, и каждой компоненты с номером j, для которого $\lambda^j{\in}B$, (любой) компонентой с номером i, для которого $\lambda^i{\in}A$. Например, если $n = 6$ и $A = \{\lambda^1, \lambda^5, \lambda^6\}$, $B = \{\lambda^2, \lambda^3\}$, то для $x^{AB} = (2, 4, 4, 3, 2, 2)$ имеем $x^{A\leftrightarrow B} = (4, 2, 2, 3, 4, 4)$.

Пусть события A и B равновероятны: $A{\sim}B$. С учетом специфики структуры векторных оценок x^{AB} и $x^{A\leftrightarrow B}$ для ЛПР они будут одинаковы по предпочтительности, что можно записать так: $x^{AB}I^{A\sim B}x^{A\leftrightarrow B}$. Следовательно, соотношение $A{\sim}B$ порождает на X рефлексивное и симметричное отношение $I^{A\sim B}$, которое включает все пары векторов вида x^{AB} и $x^{A\leftrightarrow B}$.

Пусть теперь событие A более вероятно, чем B: $A\phi B$. Рассмотрим векторную оценку x^{AB}. Пусть в ней (любая) компонента с номером i, для которой $\lambda^i{\in}A$, больше (любой) компоненты с номером j для которой $\lambda^j{\in}B$. Тогда для ЛПР векторная оценка x^{AB} будет более предпочтительна, чем $x^{A\leftrightarrow B}$, что можно записать так: $x^{AB}P^{A\not\leftrightarrow B}x^{A\leftrightarrow B}$. Следовательно, соотношение $A\phi B$ порождает на Z иррефлексивное отношение $P^{A\phi B}$, которое включает все пары векторов вида x^{AB} и $x^{A\leftrightarrow B}$, удовлетворяющие указанному условию неравенства.

Соотношение $B\phi A$ аналогично порождает на Z иррефлексивное отношение $P^{B\phi A}$.

Поскольку отношение нестрогого предпочтения R на X, которое нужно построить на основе вероятности $\propto$, должно быть квазипорядком, то его можно представить как наименьший квазипорядок на X, который включает в свой состав отношение Парето $R^{\varnothing}$ и выше определенные отношения $I^{A\sim B}$ и $P^{A\propto B}$ для всех пар событий A, B, сравнимых по $\propto$, т.е. является транзитивным замыканием объединения всех этих отношений. Определенное указанным образом отношение будем обозначать $R^{\propto}$ Таким образом, $xR^{\propto}y$ верно тогда и только тогда, когда существует цепочка

$$xR^{\pi_1}z^1,\ z^1R^{\pi_2}z^2,\ ...,\ z^tR^{\pi_{t+1}}y\ , \tag{1.4}$$

где $z^1, z^2, \ldots, z^t \in X$ и каждое R^{π_l}, $l = 1, 2, \ldots, t + 1$, есть $R^{\varnothing}$, $I^{A\sim B}$ или $P^{A\propto B}$.

Данное определение отношения $R^{\propto}$ неконструктивно. Поэтому практически важно указать эффективные методы его построения хотя бы для отдельных частных случаев.

Рассмотрим случай, когда множество оценок последствий конечно:

$$V = \{v^1, v^2, \ldots, v^n\},\ \text{где}\ v^1 < v^2 < \ldots < v^n. \tag{1.5}$$

Тогда множество векторных оценок $X = V^m$ также конечно (в нем будет $N = n^m$ элементов – векторных оценок x). Бинарное отношение ρ на X можно задать квадратной булевой матрицей смежности $B(\rho) = \left\| b_{ij}(\rho) \right\|$ порядка N, где

$$\left\| b_{ij}(\rho) = \begin{cases} 1 \text{ при } (x^i, x^j) \in \rho, \\ 0 \text{ при } (x^i, x^j) \notin \rho. \end{cases} \right\|$$

Для булевых матриц определены операции: транспонирование ^{T}, сложение $+$, поэлементное умножение $*$, умножение $\times$, образование асимметричной разности $\backslash$:

$$\|a_{ij}\| \backslash \|b_{ij}\| = \|c_{ij}\|, \text{ где } \|c_{ij}(\rho)\| = \begin{cases} 1 & \text{при } a_{ij} = 1 \text{ и } b_{ij} = 0, \\ 0 & \text{в остальных случаях.} \end{cases}$$

Операциям над бинарными отношениями – взятия обратного отношения $^{-1}$, пересечения отношений $\cap$, объединения отношений $\cup$, разности отношений $\backslash$, умножения отношений $\circ$ и построения транзитивного замыкания $\hat{\rho}$ соответствуют операции над их матрицами смежности [28]:

$$B(\rho^{-1}) = B(\rho)^{\mathrm{T}},\ B(\rho^1 \cap \rho^2) = B(\rho^1) \circ B(\rho^2),\ B(\rho^1 \cup \rho^2) = B(\rho^1) + B(\rho^2),$$

$$B(\rho^1 \backslash \rho^2) = B(\rho^1) \backslash B(\rho^2),\ B(\rho^1 \circ \rho^2) = B(\rho^1) \times B(\rho^2),$$

$$B(\hat{\rho}) = B(\rho) + B^2(\rho) + \ldots + B^n(\rho).$$

Поэтому $B(\mathrm{Sym}\ \rho) = B(\rho) \circ B(\rho)^{\mathrm{T}}$, $B(\mathrm{As}\ \rho) = B(\rho) \backslash B(\mathrm{Sym}\ \rho)$. Для построения транзитивного замыкания существуют эффективные алгоритмы [29]. Заметим еще, что

$$\rho^1 \subseteq \rho^2 \Leftrightarrow B(\rho^1) \leqq B(\rho^2),\ \rho^1 \subset \rho^2 \Leftrightarrow B(\rho^1) \leq B(\rho^2).$$

Таким образом, с помощью матричных операций можно построить отношение $R^{\propto}$. При использовании современной вычислительной техники этот путь реализуем при «не очень большой» размерности задачи («не очень больших» числах m и n). Однако этот путь, даже если его удается реализовать, может оказаться весьма обременительным. Поэтому практически важной является разработка эффективных методов построения отношений $R^{\propto}$, не ограниченных размерностью задач, для различных частных видов отношения вероятности $\propto$.

1.2.5 Случай порядковой вероятности

Для описания метода построения отношения $R^{\propto}$, когда вероятность является частично порядковой, введем в рассмотрение множества (где x, y – произвольные векторные оценки):

$$Z^{+}(y) = \{x \in X \mid \text{ существует цепочка (1.4), в которой нет } R^{\varnothing}\},$$

$$Z^{-}(x) = \{y \in X \mid \text{ существует цепочка (1.4), в которой нет } R^{\varnothing}\},$$

$$Z^{=}(x) = \{y \in X \mid \text{ существует цепочка (1.4), в которой все } R^{\pi_t} \text{ суть } I^{A\sim B}\} \cup \{x\}.$$

Понятно, что каждое из этих множеств содержит не более чем $m!$ элементов.

Теорема 1.3. Для частичной порядковой вероятности $\propto$ справедливы утверждения:

1) $xR^{\propto}y$ верно тогда и только тогда, когда существует $z \in Z^{+}(y)$, такой, что $x \geqq z$;

2) $xR^{\propto}y$ верно тогда и только тогда, когда существует $z \in Z^{-}(x)$, такой, что $z \geqq y$;

3) $xI^{\sim}y$ верно тогда и только тогда, когда $y \in Z^{=}(x)$, или, что равносильно, когда $x \in Z^{=}(y)$.

Утверждения 2) и 3) из этой теоремы для многокритериальных задач были доказаны в [30]. Утверждение 1) доказывается аналогично.

Пример 1.1. Пусть $m = 4$; частичная порядковая вероятность $\propto$ содержит, помимо пар вида (λ^i, λ^i), еще $\lambda^1 \not\sim \lambda^2$ и $\lambda^3 \sim \lambda^4$; $x = (2, 1, 6, 5)$, $y = (1, 2, 4, 6)$. Имеем: $Z^{+}(y) = \{(1, 2, 4, 6), (2, 1, 4, 6), (1, 2, 6, 4), (2, 1, 6, 4)\}$. Поскольку $z = (2, 1, 6, 4) \in Z^{+}(y)$ и верно $x \geqq z$, то, согласно теореме 1.3, справедливо $xR^{\propto}y$. Далее, $Z^{=}(y) = \{(1, 2, 4, 6), (1, 2, 6, 4)\}$ и $x \notin Z^{=}(y)$. Поэтому $xI^{\sim}y$ неверно. В итоге имеем $xP^{\propto}y$.

Теорема 1.3 прямо указывает на алгоритмический метод построения отношения $R^{\propto}$. Для сравнения двух векторных оценок он эффективен даже при

«не очень большом» числе n, но для построения всего отношения R на X этот метод может оказаться чрезмерно обременительным. Однако для полной порядковой вероятности, задаваемой (1.3), существуют аналитические методы построения $R^{\propto}$.

Для произвольных $x, y \in X$ введем в рассмотрение множество чисел, которые являются компонентами вектора x или y:

$$W(x, y) = \{x_1\} \cup \{x_2\} \cup \ldots \cup \{x_m\} \cup \{y_1\} \cup \{y_2\} \cup \ldots \cup \{y_m\}.$$

Число q элементов этого множества зависит от x и y. Обозначим эти числа через w^k и перенумеруем их по возрастания, так что имеем: $w^1 < w^2 < \ldots < w^q$. Пусть $h_{kt}(x)$ – число компонент вектора x с номерами i из групп M_1, M_2, …, M_t (см. (1.3)) таких, что $x_i > w^k$; аналогичный смысл имеет $h_{kt}(y)$. Введем в рассмотрение $(q-1)\times\tau$-матрицы $H(x) = \|h_{kt}(x)\|$ и $H(y) = \|h_{kt}(y)\|$.

Теорема 1.4. Для (полной) порядковой вероятности соотношение $xR^{\propto}y$ верно тогда и только тогда, когда $H(x) \geqq H(y)$; при этом $xP^{\propto}y$, когда $H(x) \geq H(y)$, и $xI^{\propto}y$, когда $H(x) = H(y)$.

Теорема 1.4 для многокритериальных задач была представлена в [31].

Если множество оценок V конечно, то его в теореме 1.4 можно использовать вместо множества $W(x, y)$.

Пример 1.2. Пусть $m = 4$; (полную) порядковую вероятность $\propto$ можно кратко представить в виде $\lambda^1 \not\propto \lambda^2 \sim \lambda^3 \not\propto \lambda^4$; $x = (4, 5, 2, 1)$, $y = (1, 2, 5, 4)$. Имеем: $\tau = 3$, $M_1 = \{1\}$, $M_2 = \{2, 3\}$, $M_3 = \{4\}$; $W(x, y) = \{1, 2, 4, 5\}$, $q = 4$ и

$$H(x) = \begin{Vmatrix} 1 & 3 & 3 \\ 1 & 2 & 2 \\ 0 & 1 & 1 \end{Vmatrix}, \quad H(y) = \begin{Vmatrix} 0 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 1 & 1 \end{Vmatrix}.$$

Поскольку здесь $H(x) \geq H(y)$, то, по теореме 1.4, верно $xP^{\propto}y$.

Обозначим через $a_{\downarrow} = (a_{[1]}, a_{[2]}, \ldots, a_{[n]})$ вектор, образованный из вектора $a = (a_1, \ldots, a_n)$ упорядочением его компонент по невозрастанию: $a_{[1]} \geqslant \ldots \geqslant a_{[n]}$. Под $a^{[1,j]}$ будем понимать вектор, составленный из первых j компонент вектора

a, т.е. $a^{[1,j]} = (a_1,\ldots, a_j)$. Запись $a_{\downarrow}^{[1,j]}$ будет обозначать вектор, полученный из вектора $a^{[1,j]}$ упорядочением его компонент по невозрастанию. Например, если $a = (3, 4, 1)$, то $a^{[1,2]} = (3, 4)$ и $a_{\downarrow}^{[1,2]} = (4, 3)$.

Теорема 1.5. Для (полной) порядковой вероятности соотношение $xR^{\infty}y$ верно тогда и только тогда, когда

$$x_{\downarrow}^{[1,t]} \geqq y_{\downarrow}^{[1,t]},\ t = i_1,\ldots, i_\tau;$$

при этом $xP^{\infty}y$, когда $x_{\downarrow}^{[1,t]} \geq y_{\downarrow}^{[1,t]}$ хотя бы для одного t, и $xI^{\infty}y$, когда $x_{\downarrow}^{[1,t]} = y_{\downarrow}^{[1,t]}$ для всех $t = i_1,\ldots, i_\tau$.

Эта теорема для многокритериальных задач была представлена в [31] без доказательства. В предположении существования функции полезности она была доказана в [32], а без такого допущения – в [33].

Пример 1.3. В условиях примера 1.2:

$$x^{[1,1]} = 4,\ x^{[1,3]} = (4, 5, 2),\ x^{[1,4]} = (4, 5, 2, 1);\ x_{\downarrow}^{[1,1]} = 4,\ x_{\downarrow}^{[1,3]} = (5, 4, 2),$$

$$x_{\downarrow}^{[1,4]} = (5, 4, 2, 1);$$

$$y^{[1,1]} = 1,\ y^{[1,3]} = (1, 2, 5),\ x^{[1,4]} = (1, 2, 5, 4);\ y_{\downarrow}^{[1,1]} = 1,\ y_{\downarrow}^{[1,3]} = (5, 2, 1),$$

$$y_{\downarrow}^{[1,4]} = (5, 4, 2, 1).$$

Поскольку здесь $x_{\downarrow}^{[1,1]} > y_{\downarrow}^{[1,1]}$, $x_{\downarrow}^{[1,3]} \geq y_{\downarrow}^{[1,3]}$ и $x_{\downarrow}^{[1,4]} = y_{\downarrow}^{[1,4]}$, то, по теореме 1.5, верно $xP^{\infty}y$.

1.2.6 Отношения предпочтения и безразличия при известном отношении ЛПР к риску

Под лотереей, используя принятую в теории принятия решений при риске терминологию [34], будем понимать случайную величину с конечным числом значений из V, вероятности которых известны. Лотерея называется невырожденной, если ни одна из этих вероятностей не равна 1. Здесь примем, что множество оценок V является числовым промежутком, причем «физическая» шкала критериев не менее совершенна, чем шкала отношений.

По определению, ЛПР не склонен (соответственно, склонен, безразличен) к риску, если средний ожидаемый выигрыш любой невырожденной лотереи (т.е. получение наверняка выигрыша, равного математическому ожиданию случайного выигрыша) более предпочтителен (соответственно, менее предпочтителен, безразличен) самой лотереи (т.е. участию в ней) [34].

Для всякого $x \in X$ определяются верхний и нижний срезы отношения R через x:

$$R_{+}(x) = \{y \in X \mid yRx\}, \quad R_{-}(x) = \{y \in X \mid xRy\}.$$

Квазипорядок R называется вогнутым (соответственно, выпуклым), если верхний (соответственно, нижний) его срез через любой $x \in X$ является выпуклым множеством.

Теорема 1.4. Справедливы утверждения:

1) для не склонного к риску ЛПР (γ^1) квазипорядок R является вогнутым;

2) для склонного к риску ЛПР (γ^2) квазипорядок R является выпуклым;

3) для безразличного к риску ЛПР (γ^3) квазипорядок R является и вогнутым и выпуклым.

Эта теорема была доказана в [21],

Согласно теореме 1.4 отношение нестрогого предпочтения $R^{\propto 1}$, задаваемое на основе вероятности $\propto$ для ЛПР, не склонного к риску (γ^1), определяется как наименьший вогнутый квазипорядок, продолжающий $R^{\propto}$ на X, т.е. должно выполняться $R^{\propto} \subset R^{\propto 1}$ (это влечет $I^{\sim} \subseteq I^{\sim 1}$) и $P^{\propto} \subseteq P^{\propto 1}$. Верхние срезы отношения $R^{\propto 1}$ являются выпуклыми замыканиями соответствующих срезов отношения $R^{\propto}$. Аналогично, отношение нестрогого предпочтения $R^{\propto 2}$, задаваемое на основе вероятности $\propto$ для ЛПР, склонного к риску (γ^2), определяется как наименьший выпуклый квазипорядок, продолжающий $R^{\propto}$ на X. Нижние срезы отношения $R^{\propto 2}$ являются выпуклыми замыканиями соответствующих нижних срезов отношения $R^{\propto}$. Аналогично, отношение нестрогого предпочтения $R^{\propto 3}$,

задаваемое на основе вероятности $\propto$ для ЛПР, безразличного к риску (γ^3), определяется как наименьший одновременно вогнутый и выпуклый квазипорядок, продолжающий $R^{\propto}$ на X. Верхние и нижние срезы отношения $R^{\propto 3}$ являются выпуклыми замыканиями соответствующих верхних и нижних срезов отношения $R^{\propto}$.

Данное определение отношения $R^{\propto}$ неконструктивно. Однако для случая порядковой вероятности $\propto$ можно предложить эффективный метод построения этого отношения.

Введем обозначения $\overline{R}_+(x)$, $\overline{R}_-(x)$, $\overline{Z}^+(x)$ и $\overline{Z}^-(x)$ для выпуклых замыканий множеств $R_+(x)$, $R_-(x)$, $Z^+(x)$ и $Z^-(x)$ соответственно. Заметим, что $xR^{\propto 1}y$ (соответственно, $xR^{\propto 2}y$; $xR^{\propto 3}y$), верно тогда и только тогда, когда справедливо $x \in \overline{R}_+(y)$ (соответственно, $y \in \overline{R}_-(x)$; $x \in \overline{R}_+(y)$ или $y \in \overline{R}_-(x)$).

Теорема 1.5. Для частичной порядковой вероятности $\propto$ справедливы утверждения:

1) $xR^{\propto 1}y$ верно тогда и только тогда, когда существует $z \in \overline{Z}^+(y)$, такой, что $x \geqq z$;

2) $xR^{\propto 2}y$ верно тогда и только тогда, когда существует $z \in \overline{Z}^-(x)$, такой, что $z \geqq y$;

3) $xR^{\propto 3}y$ верно тогда и только тогда, когда существует $z \in \overline{Z}^+(y)$, такой, что $x \geqq z$, или когда существует $z \in \overline{Z}^-(x)$, такой, что $z \geqq y$.

Теорема 1.5 указывает эффективный метод проверки справедливости соотношений $xR^{\propto 1}y$, $xR^{\propto 2}y$ и $xR^{\propto 3}y$. Например, для проверки соотношения $xR^{\propto 1}y$ нужно выяснить, существует ли в выпуклом замыкании $\overline{Z}^+(y)$ множества $Z^+(y)$ вектор z такой, что верно $z \geqq y$. А так как множество $Z^+(y)$ содержит конечное число векторов, то последняя задача сводится к проверке совместности

системы линейных неравенств и равенства, которую можно осуществить, используя любую компьютерную систему решения задач линейного программирования.

Пример 1.4. Пусть $m = 4$; частичная порядковая вероятность $\propto$ содержит, помимо пар вида (λ^i, λ^i), еще $\lambda^1 \not\sim \lambda^2$ и $\lambda^3 \sim \lambda^4$; $x = (3, 4, 3, 3)$, $y = (1, 5, 2, 4)$. Имеем: $Z^+(y) = \{(1, 5, 2, 4), (5, 1, 2, 4), (1, 5, 4, 2), (5, 1, 4, 2)\}$. При помощи теоремы 1.3 нетрудно проверить, что $xR^{\propto}y$ неверно. Для проверки соотношения $xR^{\propto 1}y$, согласно утверждению 1 теоремы 1.5, составим систему неравенств:

$$3 \geqslant \alpha_1 + 5\alpha_2 + \alpha_3 + 5\alpha_4,$$

$$4 \geqslant 5\alpha_1 + \alpha_2 + 5\alpha_3 + \alpha_4,$$

$$3 \geqslant 2\alpha_1 + 2\alpha_2 + 4\alpha_3 + 2\alpha_4,$$

$$3 \geqslant 4\alpha_1 + 4\alpha_2 + 2\alpha_3 + 2\alpha_4,$$

$$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 1,$$

$$\alpha_1 \geqslant 0,\ \alpha_2 \geqslant 0,\ \alpha_3 \geqslant 0,\ \alpha_4 \geqslant 0.$$

Эта система совместна: например, подходят значения $\alpha_1 = 0{,}25$; $\alpha_2 = 0{,}25$; $\alpha_3 = 0{,}5$; $\alpha_4 = 0$. Следовательно, верно $xR^{\propto 1}y$. А поскольку одно из четырех записанных выше неравенств выполняется как строгое, то верно $xP^{\propto 1}y$. Заметим, что в роли z здесь выступил вектор

$$(2, 4, 3, 3) = 0{,}25\cdot(1, 5, 2, 4) + 0{,}25\cdot(5, 1, 2, 4) + 0{,}5\cdot(1, 5, 4, 2) + 0\cdot(5, 1, 4, 2).$$

1.2.7 Отношения предпочтения и безразличия при шкале оценок первой порядковой метрики

Везде ранее полагалось, что шкала оценок V в модели (1.1) для предпочтений порядковая. Здесь мы рассмотрим случай, когда эта шкала конечна (причем $n > 2$) и является шкалой первой порядковой метрики [12], т.е. упорядочены по предпочтительности не только сами оценки v^j, но и их разности.

Ограничимся рассмотрением случая, когда эти разности вдоль V убывают (что соответствует «убыванию предельной ценности»):

$$v^2 - v^1 > v^3 - v^2 > \ldots > v^n - v^{n-1}. \qquad (1.6),$$

Далее для простоты записи положим, что $V = \{1, 2, \ldots, m\}$. Информацию о наличии шкалы первой порядковой метрики обозначим через Δ. Эта информация задает на X следующие отношения предпочтения:

$$xP^{A\sim B\&\Delta}y^{AB} \Leftrightarrow [(x = (y^{AB}|y_i + \delta, i{:}\ \lambda^i \in A; y_j - \delta, j{:}\ \lambda^j \in B, y_i + \delta \le y_j - \delta) \vee (x = (y^{AB}|y_j + \delta, j{:}\ \lambda^j \in B; y_i - \delta, i{:}\ \lambda^i \in A), y_j + \delta \le y_i - \delta)], x \in X;$$

$$xP^{A\succ B\&\Delta}y^{AB} \Leftrightarrow (x = (y^{AB}|\ y_i + \delta, i{:}\ \lambda^i \in A; y_j - \delta, j{:}\ \lambda^j \in B), y_i + \delta \le y_j - \delta), x \in X \qquad (1.7)$$

(числа δ в (1.7) – натуральные).

Отношение $R^{\propto\Delta}$ определяется как транзитивное замыкания всех отношений $I^{A\sim B}$, $P^{A\succ B}$, $P^{A\sim B\&\Delta}$, $P^{A\succ B\&\Delta}$ и $R^{\varnothing}$. Таким образом, $xR^{\propto\Delta}y$ верно тогда и только тогда, когда существует цепочка (1.4), в которой $z^1, z^2, \ldots, z^t \in X$ и каждое R^{π_l}, $l = 1, 2, \ldots, t$, есть $I^{A\sim B}$, $P^{A\not\sim B}$, $P^{A\sim B\&\Delta}$, $P^{A\succ B\&\Delta}$ или $R^{\varnothing}$.

Так как множество оценок V конечно, то отношение $R^{\propto\Delta}$ можно построить с помощью его матричного представления. Для случая (полной) порядковой вероятности $\mathcal{P}$ проверку справедливости соотношения $xR^{\propto\Delta}y$ можно провести при помощи алгоритмического метода из [35].

Если, помимо информации (1.6), известно отношение ЛПР к риску, например, что он не склонен к риску (γ^1), то соответствующе отношение нестрогого предпочтения $R^{\propto\Delta 1}$ определяется как наименьший вогнутый квазипорядок, продолжающий $R^{\propto\Delta}$ на X. Верхние срезы отношения $R^{\propto\Delta 1}$являются выпуклыми замыканиями соответствующих срезов отношения $R^{\propto\Delta}$.

1.2.8 Отношения предпочтения и безразличия при информации о сравнительной степени превосходства в вероятности

Пусть модель вероятностей П, помимо отношения вероятности $\propto$, содержит также четырехместные отношения $\triangleright$ и $\approx$, сравнивающие интенсивности превосходства в вероятности одних событий над другими (эти отношения родственны отношениям сравнительного превосходства в важности одних критериев над другими [17]).

Пусть A, B, C, D – четыре попарно несовместных события таких, что $A \propto B$ и $C \propto D$, причем превосходство в вероятности A над B больше (или же равно), чем превосходство в вероятности C над D. Последний факт будет обозначать так:

$$[A \propto B] \triangleright [C \propto D] \text{ (соответственно } [A \propto B] \approx [C \propto D]).$$

Рассмотрим векторную оценку $x^{AB,CD} = (x_1, \ldots, x_m)$, в которой все компоненты с номерами i, для которых λ^i входит в одно и то же множество из A, B, C, D, равны между собой. Под x_{i_A} (соответственно, под x_{i_B}, x_{i_C} x_{i_D}) будем понимать число, равное компоненте x_i вектора $x^{AB,CD}$ при $\lambda^i \in A$ (соответственно, при $\lambda^i \in B$, $\lambda^i \in C$, $\lambda^i \in D$). Обозначим через $x^{A \leftrightarrow B, C \leftrightarrow D}$ вектор, полученный из $x^{AB,CD}$ заменой каждой компоненты с номером i, для которого $\lambda^i \in A$, (любой) компонентой с номером x_{i_B}, и каждой компоненты с номером i, для которого $\lambda^i \in B$, (любой) компонентой с номером x_{i_A}, и аналогично для C и D. Например, если $n = 7$ и $A = \{\lambda^1\}$, $B = \{\lambda^3, \lambda^4\}$, $C = \{\lambda^5, \lambda^6\}$, $D = \{\lambda^7\}$, то для $x^{AB,CD} = (2, 3, 4, 4, 1, 1, 5)$ имеем $x_{i_A} = 2$, $x_{i_B} = 4$, $x_{i_C} = 1$, $x_{i_D} = 5$ и $x^{A \leftrightarrow B, C \leftrightarrow D} = (4, 3, 2, 2, 5, 5, 1)$.

Пусть верно $[A \propto B] \approx [C \propto D]$. Тогда, с учетом специфики векторных оценок $x^{AB,CD}$ и $x^{A \leftrightarrow B, C \leftrightarrow D}$, оказываются определенными на X отношения безразличия и предпочтения:

$x^{AB,CD} I^{[A\propto B]\approx[C\propto D]} x^{A\leftrightarrow B, C\leftrightarrow D}$ верно тогда и только тогда, когда $x_{i_A} = x_{i_D} \neq x_{i_B} = x_{i_C}$;

$x^{AB,CD} P^{[A\propto B]\approx[C\propto D]} x^{A\leftrightarrow B, C\leftrightarrow D}$ верно тогда и только тогда, когда $x_{i_B} \geqslant x_{i_C} > x_{i_D} \geqslant x_{i_{BA}}$, причем хотя бы одно из двух нестрогих неравенств выполняется как строгое.

Пусть верно $[A\propto B] \triangleright [C\propto D]$. Тогда, с учетом специфики векторных оценок $x^{AB,CD}$ и $x^{A\leftrightarrow B, C\leftrightarrow D}$, оказывается определенным на X отношение предпочтения:

$$x^{AB,CD} P^{[A\propto B]\triangleright[C\propto D]} x^{A\leftrightarrow B, C\leftrightarrow D}$$

верно тогда и только тогда, когда $x_{i_B} \geqslant x_{i_C} > x_{i_D} \geqslant x_{i_{BA}}$.

Отношение нестрогого предпочтения, порождаемое на X вероятностью $\propto$, дополненной сведениями о равенстве или превосходстве в вероятности одних событий над другими, определяется как транзитивное замыкание объединения $R^{\varnothing}$ и отношений всех трех введенных видов. Иначе говоря, в цепочке (1.4) каждое R^{π_l}, $l = 1, 2, \ldots, t + 1$, есть $R^{\varnothing}$, $I^{A\sim B}$, $P^{A\propto B}$, $I^{[A\propto B]\approx[C\propto D]}$, $P^{[A\propto B]\approx[C\propto D]}$ или $P^{[A\propto B]\triangleright[C\propto D]}$.

В случае, когда множество оценок V конечно, это отношение можно построить при помощи его матричного представления.

1.2.9 Некоторые выводы по исследованию

1) Методы, использующие качественную информацию о предпочтениях и вероятностях, позволяют сравнить по предпочтительности не все стратегии и потому часто могут не привести к получению решения требуемого типа (например, выбрать одну наилучшую стратегию). Поэтому, согласно итеративному подходу к анализу задач принятия решений [36], следует использовать дополнительную, более сильную, но и более сложную информацию. При

невозможности получения дополнительной информации возможно использование согласительных, или суррогатных решений [37].

2) В работе представлен арсенал различных по эффективности методов анализа задач принятия решений при неопределенности, использующих нечисловые оценки предпочтений и вероятностей. К сожалению, он явно недостаточен для удовлетворения запросов практики. Поэтому актуальной является проблема разработки эффективных методов подобного рода для различных сочетаний видов информации о предпочтениях и вероятностях как в рамках теории принятия решений при неопределенности, так и в рамках теории принятия многокритериальных решений.

3) В теории важности критериев разработаны также методы анализа решений, использующие количественные оценки важности и опирающиеся на строгое определение понятия превосходства в важности одних критериев над другими в некоторое число раз [38, 13, 39, 40]. Рассмотрение вопроса о приложении этих методов к задачам анализа решений при неопределенности – цель дальнейших исследований.

4) Поставленные перед началом исследований задачи выполнены полностью.

Результаты проведенных исследований нашли отражение в статье [41].

1.3 Аксиоматический анализ способов выбора оптимальных вариантов

1.3.1 Выбор, предпочтения, полезность

Формальным описанием процесса выбора является задание функции выбора C. Её аргументом является предъявление X – множество вариантов, предъявленных (физически доступных) для выбора. Примерами предъявления являются меню в ресторане и список кандидатов на общественную должность в бюллетене для голосования. Значениями предъявления как переменной величины являются непустые подмножества общего множества альтернатив A,

то есть, множества всех мыслимых вариантов выбора. Примером общего множества альтернатив является список всех лиц, имеющих право занять общественную должность. Значением функции выбора $C(X)$ является некоторое подмножество предъявления, $C(X)\subseteq X$, интерпретируемое либо как результат выбора, если множество содержит только одну альтернативу, либо как множество альтернатив, равноценных для индивида, если таковых несколько.[2]

Такое описание выбора опирается на предположение, что из одного и того же предъявления выбирающий всегда выбирает одни и те же варианты. Это предположение можно назвать нулевой аксиомой выбора. На практике оно выполняется не всегда и не строго, поэтому описание выбора с помощью функций выбора не является точным и универсальным. Однако это весьма полезное приближение, позволяющее анализировать выбор строго математически.

Субъект выбора считается действующим рационально, если его функция выбора удовлетворяет следующим двум условиям:

– непустота выбора: $\forall X\subseteq A,\ C(X)\neq\varnothing$;

– независимость от посторонних альтернатив по Нэшу [42], называемая также Аксиомой выбора по Эрроу, Arrow Choice Axiom, ACA:

$$\forall X\subseteq A,\ \forall Y\subseteq X,\ C(X)\cap Y\neq\varnothing \Rightarrow C(Y)=C(X)\cap Y.$$

Смысл этих условий следующий. Поскольку рациональный субъект обычно понимается как субъект сравнивающий, оценивающий и действующий целенаправленно в соответствии со своими оценками, то невыделение никакого подмножества из предъявления означает признание всех альтернатив (субъективно) равноценными с точки зрения достижения цели, породившей

[2] В случае $|C(X)|>1$ полагают, что выбирающему безразлично, какая именно альтернатива из содержащихся в $C(X)$ будет, в итоге, выбрана, поэтому окончательный выбор варианта уже не является выбором данного субъекта, а определяется иными факторами или случаен. В литературе функциями выбора часто именуют только $C(X)$: $\forall X\subseteq A,\ |C(X)|=1$, тогда как иные $C(X)$ называют, не функциями, а соответствиями выбора. Для целей настоящей работы данное различение не является принципиальным и не отражено в терминах.

ситуацию выбора и являющейся его ratio. Пустой выбор в такой ситуации эквивалентен выбору всего предъявления, поэтому без потери общности обе ситуации можно отождествить. Второе условие есть предположение, что удаление из меню неоптимальных (невыбираемых) и части оптимальных (выбираемых из X) вариантов не влияет на оценку остающихся в меню альтернатив как оптимальных или неоптимальных – остающиеся доступными для выбора варианты, которые выбирающих считал оптимумом, он продолжает считать оптимумом, а варианты, которые он ранее отвергал, он продолжает отвергать. Например, если из меню, состоявшего из банана, яблока и апельсина, можно было выбрать только один фрукт, и потребитель выбирал яблоко, то предполагается, что исключение из меню банана не должно приводить к изменению потребительского выбора, если потребитель рационален.

Ранжированием альтернатив называется слабый порядок, то есть бинарное отношение R на A, $R \subseteq A \times A$, удовлетворяющее двум аксиомам:

– полнота: все альтернативы сравнимы, $\forall x, y \in A, xRy \vee yRx$;

– транзитивность: $\forall x, y, z \in A, (xRy \wedge yRz) \Rightarrow xRz$.

Имеет место следующая характеристическая теорема.

Теорема 1.6. Функция выбора $C(X)$ удовлетворяет аксиомам Непустоты и Независимости от посторонних альтернатив по Нэшу тогда и только тогда, когда существует ранжирование R альтернатив из A, такое что $C(X)=\mathrm{MAX}(R|_X)$ для любого предъявления X. Такое ранжирование единственно.

Здесь и далее $R|_X=R \cap X \times X$ – это *сужение* отношения R ($R \subseteq A \times A$) на подмножество X ($X \subseteq A$). $MAX(R|_X)=\{a \in X \mid \forall b \in X, bRa \Rightarrow aRb\}$ – это множество максимальных элементов подотношения $R|_X$.

Таким образом, рациональный субъект есть субъект максимизирующий, а его выбор рационализируется (объясняется и предсказывается) его предпочтениями, математической репрезентацией которых является бинарное отношение R.

В случае, когда A конечно или счётно, любое ранжирование R альтернатив из A, а в случае, когда A имеет мощность континуума, любое непрерывное ранжирование R может, в свою очередь, быть репрезентировано функцией $u(x)$: $A \to \mathbb{R}$, такой что $u(y) \geq u(x) \Leftrightarrow yRx$. Число $u_0 = u(x)$ называется полезностью, которую рациональный субъект с предпочтениями R извлекает из альтернативы x, если она становится результатом выбора. Соответственно, $u(x)$ называется его функцией полезности.

Какой смысл следует придавать функции $u(x)$? Если полезность или ненаблюдаема, или неизмеряема, или не имеет строгого определения, или попросту фиктивна, тогда $u(x)$ есть не более чем удобная математическая репрезентация R. Если полезность наблюдаема и измерима с помощью какой-либо ординальной (порядковой) или кардинальной (количественной) шкалы, тогда $u(x)$ есть результат её измерения или оценки. Когда ранжирование основано на результатах такой оценки оно называется рейтингом.

В общем случае функция полезности определяется неоднозначно. Если полезность выбирающего неизмерима или измерима лишь ординально, тогда возможно заменить $u(x)$ на $u'(x) = \varphi(u(x))$, где преобразование φ – это произвольная строго возрастающая функция φ: $\mathbb{R} \to \mathbb{R}$. Аналогично, если полезность измерима количественно, но ни точка отсчёта, ни единица измерения шкалы не фиксированы, тогда возможно заменить $u(x)$ на $u'(x) = a \cdot u(x) + b$, где b – произвольное, a – произвольное положительное действительные числа. Такое преобразование называют аффинным.

Если субъектов выбора несколько, а объект выбора один и тот же,[3] и если не существует общей для всех субъектов шкалы измерения полезности результата выбора, кардинальной или ординальной, то их функции полезности могут быть трансформированы независимо друг от друга. Это означает, что полезности, которые субъекты извлекают из одинаковых благ или из общего

[3] Таким объектом могут являться как идентичные индивидуально потребляемые блага, так и состояние мира общее для всех субъектов выбора, становящееся результатом их коллективного выбора.

состояния мира, несравнимы. В результате этой несравнимости все утверждения типа “положение Анны лучше положения Бориса” будут непроверяемыми.

– ординальная несравнимость полезностей:

набор функций $\{u_k(x)\}$, $k=1\div n$, во всех вычислениях можно заменить набором $\{u'_k(x)\}$, $k=1\div n$, $u'_k(x)=\varphi_k(u_k(x))$, где преобразования φ_k – это произвольные (и, вообще говоря, различные) строго возрастающие функции φ_k: $\mathbb{R} \to \mathbb{R}$.

– ординальная сравнимость полезностей:

набор функций $\{u_k(x)\}$, $k=1\div n$, во всех вычислениях можно заменить набором $\{u'_k(x)\}$, $k=1\div n$, $u'_k(x)=\varphi(u_k(x))$, где преобразование φ – это произвольная, но одна и та же для всех субъектов k, строго возрастающая функция φ: $\mathbb{R} \to \mathbb{R}$.

Ординальная сравнимость полезностей подразумевает существование общей для всех порядковой шкалы измерения индивидуальной полезности. Эта шкала создаёт возможность порядковых сравнений благосостояния различных лиц.

– кардинальная несравнимость полезностей:

набор функций $\{u_k(x)\}$, $k=1\div n$, во всех вычислениях можно заменить набором $\{u'_k(x)\}$, $k=1\div n$, $u'_k(x)=a_k u(x)+b_k$, где $\{a_k\}$ и $\{b_k\}$ – это наборы из n произвольных ($a_k>0$ для всех k) действительных чисел.

– кардинальная сравнимость полезностей:

набор функций $\{u_k(x)\}$, $k=1\div n$, во всех вычислениях можно заменить набором $\{u'_k(x)\}$, $k=1\div n$, $u'_k(x)=au(x)+b$, где a и b – два произвольных ($a>0$) действительных числа.

Кардинальная сравнимость полезностей подразумевает существование общей для всех количественной шкалы измерения индивидуальной полезности. Эта шкала создаёт возможность как порядковых, так и количественных сравнений благосостояния различных лиц.

1.3.2 Аксиомы коллективного выбора

Необходимость сравнения благосостояния индивидов, как уже было сказано выше, возникает, когда выбор является коллективным. Математически ситуация коллективного выбора описывается следующим образом:

A – общее множество альтернатив (возможные состояния мира, общие для всех участников процесса принятия коллективных решений);

X – предъявление: $X \subseteq A \wedge X \neq \varnothing$;

N – коллектив (например, совет директоров, избирательный округ, комитет экспертов);

$u_k(x)$ – полезность альтернативы $x \in A$ для избирателя $k \in N$;

$U = \{u_k(x) \mid k \in N\}$ – профиль функций полезности.

Проблема оптимального коллективного выбора состоит в том, чтобы из каких-либо соображений оценки коллективного блага определить функцию оптимального коллективного выбора $SC(X, U) = SC(X, U)$. Однако, часто эту проблему ставят более сильно – для каждого предъявления требуют построить ранжирование его альтернатив $R(X, U)$. В этом случае, $SC(X, U) = \mathrm{MAX}(R(X, U))$. $R(X, U)$ называют функционалом коллективного благосостояния [43].

По определению, $R(X, U)$ удовлетворяет следующим аксиомам.

– полнота: $\forall X \subseteq A, \forall x, y \in X, xR(X, U)y \vee yR(X, U)x$;

– транзитивность: $\forall X \subseteq A, \forall x, y, z \in X, (xR(X, U)y \wedge yR(X, U)z) \Rightarrow xR(X, U)z$.

Поскольку в большинстве практических случаев множество A слишком велико и трудноопределимо, то узнать U полностью или слишком дорого, или вовсе невозможно. Поэтому на практике для принятия коллективных решений

используются только функционалы, требующие знания лишь небольшой части U, а именно только полезностей альтернатив предъявления.

– Локальность: $R(X, U)=R(U|_X)$.

Очевидно, что идея определить функционал $R(X, U)$ возникла из желания определить коллективные предпочтения и оптимальный коллективный выбор как выбор максимальных элементов отношения, эти предпочтения репрезентирующего. Чтобы это было возможно, локальный функционал $R(U|_X)$ должен удовлетворять следующей аксиоме.

– Независимость от посторонних полезностей: $\forall X\subseteq A$, $R(U)|_X=R(U|_X)$.

В том и только в том случае, когда $R(U|_X)$ удовлетворяет вышенаписанной аксиоме, мы можем гарантировать, что $SC(X, U)=\mathrm{MAX}(R(U|_X))$ будет удовлетворять аксиоме Независимости от посторонних альтернатив по Нэшу. Роль коллективных предпочтений R при этом будет играть ранжирование $R(U)$ множества A.

Какие ещё условия следует наложить на функционал $R(X, U)$? Поскольку часто заранее неизвестно каковы функции полезности отдельных избирателей, разумно определить $R(X, U)$ для всех мыслимых профилей U.

– Неограниченность области определения: правило ранжирования можно применить всегда, то есть при всех U.

Единственным принципом сравнения коллективного благосостояния в двух разных состояниях является принцип Парето. Сформулируем его в ослабленной форме как принцип единогласия – если для каждого члена общества альтернатива x не хуже (лучше) чем y, то и для всего общества x должно быть не хуже (лучше) чем y.

– Единогласие: $\forall X\subseteq A$, $\forall x, y\in X$,

$$(\forall k\in N, u_k(x)\geq u_k(y)) \Rightarrow xR(X, U)y) \wedge (\forall k\in N, u_k(x)>u_k(y)) \Rightarrow xP(X, U)y).$$

Здесь P обозначена ассиметричная часть R: $(x, y) \in P \Leftrightarrow((x, y) \in R \wedge (y, x) \notin R)$. Последнее требование, как и аксиома Неограниченности области определения, есть следствие нашего незнания. Во многих практически значимых

случаях невозможно или строго определить, или наблюдать, или измерять, или сравнивать индивидуальные полезности. Тогда мы, по необходимости, должны предполагать ординальную несравнимость индивидуальных полезностей (см. выше). Поэтому функционал также должен удовлетворять следующей аксиоме.

– Ординальная несравнимость. Пусть $U|_X=\{u_k(x)\}$, $U'|_X=\{u'_k(x)\}$, $x\in X$, $k=1\div n$. Если $\forall x\in X$, $u'_k(x)=\varphi_k(u_k(x))$, где φ_k – произвольные строго возрастающие функции, то $R(U|_X)=R(U'|_X)$.

То есть удовлетворяющий данному свойству функционал инвариантен относительно произвольной трансформации полезностей, не меняющей индивидуальные предпочтения.

1.3.3 Теоремы о невозможности выбора

Фактически, первым результатом, полученным в теории коллективного выбора, была знаменитая теорема Эрроу. Переформулируем её для функционалов.

Теорема 1.7 (теорема Эрроу, [44]). Если $|A|>2$, и функционал общественного благосостояния $R(X, U)$ удовлетворяет всем вышеперечисленным условиям, а именно аксиомам Полноты, Транзитивности, Локальности, Независимости от посторонних полезностей, Неограниченности области определения, Единогласия и Ординальной несравнимости, то существует диктатор d, то есть индивид, чьи предпочтения R_d совпадают с коллективными предпочтениями $R=R(A, U)$, $\exists d\in N$: $R=R_d$.

Если рассматривать функционал $R(X, U)$ как правило принятия решений, то теорема Эрроу звучит как запрет на демократические формы рационального коллективного выбора. Однако невозможность недиктаторских правил зависит от выполнения всех условий теоремы. Поэтому дальнейшее развитие теории выбора пошло по пути отказа от какого-либо из этих условий. Рассмотрим возможности, открывающиеся, на этом пути.

Начать, однако, следует с невозможностей. От каких аксиом отказ невозможен или бессмыслен? Перечислим их ещё раз:

– полнота;

– транзитивность;

– локальность;

– независимость от посторонних полезностей;

– неограниченность области определения;

– единогласие;

– ординальная несравнимость.

Во-первых, не получится отказаться от Локальности, так как использовать нелокальные процедуры практически невозможно, ибо невозможно собрать всю необходимую информацию о множестве А и о полных предпочтениях индивидов.

Во-вторых, бессмысленно отказываться от Полноты. Отказ от этой аксиомы означает лишь то, что мы допускаем пустой выбор из двух вариантов, то есть различаем случай невыбора и выбора всех вариантов. Как уже было отмечено выше, для такого различения нет достаточных оснований. Оба этих случая означают безразличие, и как следствие бездействие субъекта при определении варианта, который в итоге будет выбран.

В-третьих, отказ от принципа Единогласия, при сохранении всех остальных требований приводит к аналогу теоремы Эрроу.

Теорема 1.8 (теорема Вильсона [45]). Если $|A|>2$, и функционал общественного благосостояния $R(X, U)$ удовлетворяет аксиомам Полноты, Транзитивности, Локальности, Независимости от посторонних полезностей, Неограниченности области определения и Ординальной несравнимости, то в коллективе есть или диктатор, $\exists d \in N$: $R=R_d$, или антидиктатор, $\exists d \in N$: $\forall x, y \in A$, $xRy \Leftrightarrow yR_dx$.

Наконец, смягчение условия Ординальной несравнимости до несравнимости кардинальной, то есть допущение возможности количественного сравнения полезности двух вариантов одним индивидом при сохранении невозможности сравнения полезности одного варианта для двух индивидов, не может избавить нас от диктатора.

Теорема 1.9 [46]. Если удовлетворены аксиомы Полноты, Транзитивности, Локальности, Независимости от посторонних полезностей, Неограниченности области определения и Единогласия, то аксиома Кардинальной несравнимости логически эквивалентна аксиоме Ординальной несравнимости.

Из данной теоремы и теоремы Эрроу непосредственно следует наличие диктатора и в случае замены ординального принципа его кардинальным аналогом.

1.3.4 Отказ от аксиомы Транзитивности

Если нас интересует не ранжирование альтернатив предъявления, а оптимальный коллективный выбор, то нетривиальные возможности открывает отказ от Транзитивности. Примером такого функционала является отношение Парето R_{Pareto}, определяемое через свою ассиметричную часть Q с помощью (сильного) принципа Парето.

– Принцип Парето

$$\forall X \subseteq A,\ \forall x, y \in X,\ ((\forall k \in N,\ u_k(x) \geq u_k(y)) \wedge (\exists k \in N,\ u_k(x) > u_k(y)) \Rightarrow xQy.$$

Отношение Парето удовлетворяет всем аксиомам, кроме Транзитивности, причём в Q никогда нет циклов. В силу последнего обстоятельства оптимальный коллективный выбор, определяемый как $\text{MAX}(Q|_X)$, никогда не пуст. Отношение Парето также удовлетворяет следующей важной аксиоме.

– Монотонность. Пусть полезность одного избирателя i от одной альтернативы x увеличивается, а все остальные полезности остаются неизменными,

$$u'_i(x) > u_i(x) \wedge \forall y \in X \backslash \{x\},\ u'_i(x) = u_i(x) \wedge \forall k \in N \backslash \{i\},\ \forall y \in X,\ u'_k(x) = u_k(x)),$$

тогда коллективное мнение об x (репрезентируемое отношением $R(U|_X)$) не должно ухудшаться,

$\forall y \in X \setminus \{x\}$, $(xR(U|_X)y \Rightarrow xR(U'|_X)y) \wedge (xP(U|_X)y \Rightarrow xP(U'|_X)y)$.

Проблема с отношением Парето заключена в том, что его ассиметричная часть слишком бедна, поэтому множество вариантов оптимальных по Парето обычно слишком велико. Можно ужесточить условие монотонности требованием использовать всю имеющуюся информацию для того, чтобы различить альтернативы при парном сравнении.

– Строгая Монотонность. Пусть имеет место Монотонность. Пусть $u'_i(x)>u_i(x) \wedge \forall y \in X \setminus \{x\}$, $u'_i(x)=u_i(x) \wedge \forall k \in N \setminus \{i\}$, $\forall y \in X$, $u'_k(x)=u_k(x)) \wedge \exists z \in X$: $(u_i(x)<u_i(z) \wedge u'_i(x)=u'_i(z)) \vee (u_i(x)=u_i(z) \wedge u'_i(x)>u'_i(z))$, тогда $xR(U|_X)z \Rightarrow xP(U'|_X)z$.

Данное требование также известно под названием Положительной чувствительности.

Обобщением теоремы Мэя [47] является утверждение, что функционал, удовлетворяющий аксиомам Полноты, Локальности, Независимости от посторонних полезностей, Неограниченности области определения, Ординальной несравнимости, Строгой Монотонности и принципу Парето есть сравнение альтернатив попарно с помощью голосования на основании правила простого большинства голосов.

Таких правил бесконечно много, так как голосование может быть взвешенным, то есть голоса избирателей могут иметь неравный (но обязательно ненулевой) вес.

Правило простого большинства взвешенных голосов:

пусть $N(xPy)=\{k \in N \mid u_k(x)>u_k(y)\}$, $N(yPx)=\{k \in N \mid u_k(y)>u_k(x)\}$, v_k – вес голоса избирателя k, $\forall k \in N$, $v_k > 0$, тогда $xP_M y \Leftrightarrow \sum_{k \in N(xPy)} v_k > \sum_{k \in N(yPx)} v_k$

Асимметричное отношение P_M, получающееся с помощью правила большинства и играющее роль строгих коллективных предпочтений, называется мажоритарным. На практике в качестве весов v_k используются только рациональные неотрицательные числа. Любой же конечный набор таких чисел всегда можно, пропорционально увеличив, заменить набором чисел натуральных. Соответственно, если у избирателя k число голосов равно m_k, то данного избирателя можно заменить на m_k избирателей с одним голосом и теми же предпочтениями, что и у k. Поскольку ограничений на число избирателей нет, без потери общности можно считать, что есть только одно правило голосования, а именно правило простого большинства избирателей.

Правило простого большинства избирателей: $xP_My \Leftrightarrow |N(xPy)| > |N(yPx)|$.

Это правило, очевидно, не является диктаторским. Более того, оно удовлетворяет следующей аксиоме.

– Анонимность (равноценность голосов или равноправие избирателей): любая перестановка π(k) функций полезности в профиле U не меняет R(U): U'={$u_{\pi(k)}$(x)} ⇒ R(U')=R(U).

Однако ценой применения правила большинства, как известно, становится нарушение аксиомы Непустоты выбора, являющееся следствием возможности (в силу нарушения аксиомы Транзитивности) циклов в мажоритарном отношении: $|X|>2 \wedge |N|>2 \Rightarrow \exists U$: $MAX(P_M(U))=\varnothing$. Этот результат известен как парадокс Кондорсе [48].

В силу нарушения аксиомы Транзитивности функционал $R_M(U)$[4] более нельзя назвать функционалом общественного благосостояния, а задача ранжирования альтернатив предъявления остаётся нерешённой. У нас осталось три возможности сохранить Транзитивность: 1) отказаться от аксиомы Неограниченности области определения, 2) отказаться от аксиомы Независимости от

[4] R_M – это нестрогие коллективные предпочтения, однозначно восстанавливаемые по своей ассиметричной части P_M.

посторонних полезностей, 3) отказаться от аксиомы Ординальной несравнимости. Последовательно рассмотрим эти варианты.

1.3.5 Отказ от аксиомы Неограниченности области определения

Циклы Кондорсе в строгих коллективных предпочтениях типа $xPyPzPx$ возникают как следствие наличия трёх альтернатив x, y, и z и трёх функций полезности таких, что $u_1(x)>u_1(y)>u_1(z)$, $u_2(y)>u_2(z)>u_2(x)$, $u_3(z)>u_3(x)>u_3(y)$. В этом случае говорят, что альтернативы x, y, и z образуют тройку Кондорсе. Если ограничить множество допустимых профилей полезностей так, чтобы троек Кондорсе не было, то в мажоритарном отношении никогда не будет циклов, что решает проблему пустоты оптимального коллективного выбор. А если все индивидуальные предпочтения будут ещё и линейными порядками, то и мажоритарное отношение будет линейным порядком, что решает проблему ранжирования альтернатив любого предъявления.

Примером такого ограничения области допустимых функций полезности, является случай, когда общее множество альтернатив A является «одномерным пространством», то есть линейно упорядоченным множеством $(A, \succ)$, а функции полезности $u(x)$ удовлетворяют свойству «однопиковости» - они обладают единственной точкой максимума x^* (различной для различных функций) и монотонно убывают при удалении от неё в обе стороны данного «пространства», $\forall y, z \in A, (z \succ y \succ x^* \vee x^* \succ y \succ z) \Rightarrow u(x^*) > u(y) > u(z)$. Очевидно, что на любом конечном предъявлении $X \subseteq A$, $|X|<\infty$, свойство «однопиковости» сохраняется. В этом случае, у мажоритарного отношения P_M всегда есть один и, что особенно удобно, только один максимальный элемент m, $MAX(P_M|_X)=\{m\}$, - это медиана наилучших (в данном предъявлении) альтернатив всех избирателей, $m = \mathrm{median}\{\mathrm{argmax}_{x\in X}(u_k(x))|k \in N\}$, что составляет утверждение теоремы о медианном избирателе [49].

1.3.6 Отказ от аксиомы Независимости от посторонних полезностей

Отказ от этой аксиомы не желателен, ибо приводит к зависимости коллективных оценок не только от самих сравниваемых альтернатив, но и от контекста сравнения, то есть предъявления. Однако на этом пути открывается множество возможностей для построения функционалов общественного благосостояния. Рассмотрим примеры функционалов, основанных на мажоритарном отношении. По аксиоматическим соображениям, изложенным выше мажоритарное отношение – это единственный кандидат на роль коллективных предпочтений, которые могут оказаться нетранзитивными. Функционал коллективного благосостояния предлагается строить так. Если мажоритарное отношения в данном случае (для данного предъявления и данного профиля полезностей) является ранжированием, то оно есть значение функционала. Если мажоритарное отношение нетранзитивно, то его следует «исправить» минимальным (в некотором смысле) образом. Поскольку способов «исправления» (устранения нетранзитивностей) много, то возникает множество различных функционалов.

1.3.6.1 Правило Коупланда

Примером такого функционала является правило Коупланда [50]. Зная предъявление X, вычислим нужную часть мажоритарного отношения P_M. Отношение P_M можно рассматривать как турнир, то есть состязание, в котором первое и остальные места определяются на основании результатов парных сравнений участников. Рассмотрим функцию очков $t_\alpha(x) = |\{y \in X | xP_M y\}| + \alpha \cdot |\{y \in X | \overline{xP_M y} \wedge \overline{yP_M x}\}|, 0 \leq \alpha \leq 1$. То есть, за победу в парном состязании мы даём 1 очко, за поражение 0 очков, а за ничью α очков. Искомым ранжированием является упорядочение альтернатив по числу очков $t_\alpha(x)$. В силу своей простоты правило Коупланда широко применяется для ранжирования участников спортивных турниров и определения победителя.

Основанный на правиле Коупланда и мажоритарном отношении функционал общественного благосостояния будет удовлетворять аксиомам Полноты, Транзитивности, Локальности, Неограниченности области определения, Ординальной несравнимости, Монотонности, Анонимности и принципу Парето. В силу теоремы Эрроу он не может удовлетворять Независимости от посторонних полезностей, однако он удовлетворяет более слабому её аналогу.

Конъюнкция Ординальной несравнимости и Независимости от посторонних полезностей даёт аксиому Независимости от посторонних альтернатив по Эрроу:

– независимость от посторонних альтернатив по Эрроу. Пусть предъявление X и профиль полезностей U меняются, соответственно, на X' и U'. Тогда $\forall X,\ X'\subseteq A,\ \forall x,\ y\in X\cap X',\ (\forall k\in N,\ u_k(x)=u_k(y) \Leftrightarrow u'_k(x)=u'_k(y) \wedge u_k(x)>u_k(y) \Leftrightarrow u'_k(x)>u'_k(y)) \Rightarrow (xR(U, X)y \Leftrightarrow xR(U', X')y \wedge xP(U, X)y \Leftrightarrow xP(U', X')y)$.

То есть, предъявление и полезности могут меняться, но мнение коллектива о сравнительной пользе двух альтернатив зависит только от результатов сравнения x с y каждым избирателем, и если они не меняются, то и мнение коллектива о том, лучше ли x чем y, остаётся прежним.

Правило же Коупланда удовлетворяет следующей аксиоме [51]:

– слабая независимость от посторонних альтернатив по Эрроу. Пусть профиль полезностей U меняется U', а предъявление X не меняется. Тогда $\forall X\subseteq A, \forall x, y\in X, (\forall k\in N, \forall z\in X, u_k(x)=u_k(z) \Leftrightarrow u'_k(x)=u'_k(z) \wedge u_k(x)>u_k(z) \Leftrightarrow u'_k(x)>u'_k(z) \wedge u_k(y)=u_k(z) \Leftrightarrow u'_k(y)=u'_k(z) \wedge u_k(y)>u_k(z) \Leftrightarrow u'_k(y)>u'_k(z)) \Rightarrow (xR(U, X)y \Leftrightarrow xR(U', X)y \wedge xP(U, X)y \Leftrightarrow xP(U', X)y)$.

То есть, если предъявление не меняется, мнение коллектива о сравнительной пользе двух альтернатив, определяемое с помощью правила Коупланда, не меняется, если меняются результаты парного сравнения x и y друг с другом и с остальными альтернативами предъявления.

1.3.6.2 Сортировка с помощью турнирных решений

Рассмотрим иной способ сохранить аксиому *Независимости* в слабой форме. Как мы знаем, ранжирование альтернатив тесно связано с функцией выбора. Если мы знаем функцию выбора, и она удовлетворяет аксиомам Непустоты и Независимости от посторонних альтернатив по Нэшу, то можно построить ранжирование сортировкой. Процедура сортировки устроена следующим образом. Сначала с помощью функции выбора $C(X, U)$ мы определяем множество наилучших вариантов (альтернатив первого сорта) $X_{(1)}=C(X, U)\subseteq X$. Затем мы исключаем $X_{(1)}$ из предъявления и повторяем процедуру выбора для множества $X\backslash X_{(1)}$. Множество $X_{(2)}=C(X\backslash X_{(1)}, U)=C(C(X, U), U)$ содержит альтернативы второго сорта. После конечного числа исключений (если X конечно) предъявление будет разделено на конечное число классов эквивалентности $X_{(m)}=C(X\backslash(X_{(m-1)}\cup X_{(m-2)}\cup\ldots\cup X_{(2)}\cup X_{(1)}, U)$, что и будет искомым ранжированием. Если коллективные предпочтения репрезентируются мажоритарным отношением P_M, то должно быть $C(X, U)= MAX(P_M|_X)$. Как говорилось выше, такая функция выбора удовлетворяет Независимости от посторонних альтернатив по Нэшу, но не удовлетворяет аксиоме Непустоты в силу парадокса Кондорсе. Поэтому идея состоит в том, чтобы доопределить её разумным образом в случае, когда множество максимальных элементов мажоритарного отношения пусто. Так мы приходим к определению турнирных решений.

Функция выбора $C(X, T)$, зависящая от предъявления X и турнира T (асимметричного бинарного отношения T на общем множестве альтернатив A, $T\subseteq A\times A$), через сужение $T|_X$, $C(X, T)=C(T|_X)$, турнирным решением (способом выбора победителя в круговом турнире) если она удовлетворяет следующим аксиомам:

– непустота: $\forall T\subseteq A\times A$, $\forall X\subseteq A$, $C(T|_X)\neq\varnothing$;

– нейтральность: пусть $\pi(x)$ взаимно-однозначное отображение множества X на себя, $\pi(T|_X)\subseteq X\times X$ – турнир получающийся из турнира $T|_X$ отоб-

ражением $\pi(x)$, $(\pi(x), \pi(y)) \in \pi(T|_X) \Leftrightarrow (x, y) \in T|_X$, $\pi(Y) \subseteq X$ – множество, получающееся из множества Y отображением $\pi(x)$, $\pi(x) \in \pi(Y) \Leftrightarrow x \in Y$. Тогда $C(\pi(T|_X))=\pi(C(T|_X))$ при любых X и $\pi(x)$. Иными словами, результат выбора не зависит от имён альтернатив, правило выбора обращается со всеми альтернативами одинаково.

– Принцип Кондорсе: Всегда имеет место включение $MAX(T|_X) \subseteq C(T|_X)$. А если в турнире $T|_X$ есть однозначный победитель w (победитель Кондорсе), $\forall x \in X \setminus \{w\}$, wTx, то $C(T|_X)=MAX(T|_X)=\{w\}$.

Любой функционал общественного благосостояния, основанный на сортировке с помощью мажоритарного отношения и турнирного решения будет удовлетворять аксиомам Полноты, Транзитивности, Локальности, Неограниченности области определения, Ординальной несравнимости и Анонимности.

Если функция выбора $C(P_M(U)|_X)$ удовлетворяет принципу Парето для функций выбора, $\forall X \subseteq A$, $\forall y \in X$, $(\exists x \in X: xQy) \Rightarrow y \notin C(X, U)$, то есть, если всегда имеет место включение $C(X, U) \subseteq MAX(Q|_X)=MAX(R_{\mathrm{Pareto}}|_X)$, то принципу Парето будет также удовлетворять и соответствующий функционал. Примером таких турнирных решений являются непокрытое множество [52, 53], минимальное покрывающее множество [54], двухпартийное множество [55].

Если функция выбора $C(P_M(U)|_X)$ удовлетворяет слабому принципу Парето для функций выбора, $\forall X \subseteq A$, $\forall y \in X$, $(\exists x \in X: xQy \wedge x \notin C(X, U)) \Rightarrow y \notin C(X, U)$, то функционал $R(U|_X)$ будет также удовлетворять следующей аксиоме:

– слабый принцип Парето:

$$\forall X \subseteq A,\ \forall x, y \in X,\ (\forall k \in N,\ u_k(x) \geq u_k(y)) \Rightarrow xR(U|_X)y.$$

Примером таких решений являются максимальный цикл [56, 57, 58, 59] и объединение минимальных внешнеустойчивых множеств [6, 60].

В силу теоремы Эрроу никакие турнирные решения $C(P_M(U)|_X)$ не могут удовлетворять Независимости от посторонних альтернатив по Нэшу, но есть решения, удовлетворяющие следующей аксиоме, являющейся смягчением предыдущей:

– слабая независимость от посторонних альтернатив по Нэшу (аксиома Отбрасывания): $\forall X \subseteq A$, $\forall Y \subseteq X$, $C(X) \subseteq Y \Rightarrow C(Y)=C(X)$.

Известно четыре турнирных решения, которые удовлетворяют этому условию: максимальный цикл, минимальное покрывающее множество, двухпартийное множество и объединение минимальных внешнеустойчивых множеств. Утверждение относительно последнего решения было доказано в ходе настоящего исследования.

Соответственно, функционалы, основанные на сортировке с помощью данных решений, удовлетворяют следующему слабому варианту аксиомы Независимости от посторонних альтернатив по Эрроу:

– независимость от классов посторонних альтернатив. Пусть $\{X_{(i)}=C(X\backslash(X_{(i-1)}\cup X_{(i-2)}\cup\ldots\cup X_{(2)}\cup X_{(1)}, U)\}$ – классы эквивалентности, на которые сортировка с помощью решения $C(X, U)$ делит предъявление X. Пусть число классов – m. Выберем произвольным образом из множества первых m натуральных чисел непустое подмножество E. Образуем предъявление Y, исключив из X классы с номерами, не входящими в E, $Y = \bigcup_{i\in E} X_i$. Сортировка предъявления Y с помощью данного решения разделит Y на те же самые классы $\{X_{(i)}|\ i\in E\}$, упорядоченные также, как и в первоначальном ранжировании.

1.3.7 Отказ от аксиомы Ординальной несравнимости

Наличие единой шкалы измерения полезности есть весьма благоприятное для исследователя обстоятельство. Оно даёт возможность без особого труда построить функционалы общественного благосостояния, удовлетворяющие всем прочим аксиомам. Примером такого функционала является рейтинг, строящийся с помощью какого-либо правила агрегирования индивидуальных оценок альтернатив. Здесь интересно отметить, что для этого достаточно шкалы ординальной. Если шкала полезности является ординальной, то

полезности $u_k(x)$ данной альтернативы можно упорядочить и оценить коллективную полезность, например, как медиану, минимум или максимум данного ряда оценок. Затем альтернативы можно упорядочить по полученной таким путём единой оценке. Если же шкала является кардинальной, то наиболее популярным способом построения рейтинга альтернатив является взвешенное суммирование их полезностей и упорядочение по величине суммы.

1.4 Исследование свойств профилей предпочтений

1.4.1 Введение

В данном подразделе рассматривается классическая (по Эрроу) структура теории социального выбора. Рассматриваются n порядков предпочтений множества m-элементов, удовлетворяющих определенным свойствам. Каждый порядок предпочтений представляет собой бинарное отношение быть лучше для избирателя. Элементы множества X называются альтернативами. Набор из n порядков предпочтений называется профилем предпочтений. Анонимность не предполагается. Нетривиальная перестановка различных порядков предпочтений приводит к неэквивалентному профилю предпочтений. Перестановка равных порядков предпочтений не изменяет профиль предпочтений.

Инада [61, 62] и Сен [63] ввели несколько типов доменов Кондорсе (если общество состоит из избирателей, чьи предпочтения принадлежат домену Кондорсе, то простое правило большинства подразумевает транзитивное социальное упорядочение). Целью данной работы является анализ группы сепарабельных предпочтений. Инада [61] определил свойство отделимости следующим образом. Пусть A и B – две группы альтернатив. Для каждого возможного порядка предпочтений любая альтернатива в A всегда предпочтительнее любой альтернативы в B, или любая из B всегда предпочтительнее любой из A, тогда мы называем множество альтернатив разделимыми на две группы. Свойство группового разделения подразумевает разбиение каждого набора

альтернатив на две части таким образом, что каждый избиратель предпочитает каждую альтернативу в одной части каждой альтернативе в другой части. Несмотря на общее разделение, большая часть отличается от избирателя к избирателю.

Структурированные предпочтения имеют теоретическое значение. Есть много примеров вычислительных проблем социального выбора, которые становятся легкими, когда есть структурированные предпочтения [64, 65]. В дальнейших исследовательских работах, например, [66, 67] более слабые понятия структурированных предпочтений (около однопиковых предпочтений и т. д.) были введены. Элкинд с соавторами [68] исследовали структуры клонов в предпочтениях как полезный инструмент для изучения структурированных предпочтений. В нашем исследовании применяются структуры клонов для анализа групповых предпочтений.

Баллестер и Хэрингер [69] продемонстрировали, что групповые сепарабельные предпочтения могут быть определены с помощью запрещенных последовательностей в порядке предпочтения. Существует два типа ограничений для групповых разделяемых предпочтений. Первое условие-ограниченность среды, то есть для каждой тройки альтернатив множество альтернатив, находящихся в середине хотя бы одного подотряда, имеет мощность 1 или 2. Это ограничение было впервые обнаружено Сеном [63]. Вторым ограничением является отделимость каждой пары предпочтительных ордеров. Сепарабельность означает, что один порядок предпочтений может быть получен из другого порядка предпочтений посредством сепарабельной перестановки. Перестановка разделима тогда и только тогда, когда она не содержит ни 2413, ни 3142 в качестве шаблона. Сепарабельные перестановки подсчитываются большими числами Шредера (см. различные доказательства в [70, 71, 72]). Сепарабельные перестановки имеют различные представления дерева [73, 74, 75, 76], которые применимы для связанных комбинаторных задач.

Перечислительная комбинаторика профилей предпочтений направлена на поиск количества профилей предпочтений с требуемыми свойствами. Он

применяется для вычисления соответствующих вероятностей. Предположение беспристрастной культуры утверждает, что каждый возможный порядок предпочтений имеет одинаковую вероятность возникновения.

Герлейн и Фишберн [77, 78] получили рекурсивные и точные формулы для числа профилей предпочтений с победителем Кондорсе. Лакнер и Лакнер [79] получили асимптотическую формулу числа однопиковых профилей предпочтений. Чен и Финнендал [80] обнаружили количество одноконечных нарциссических и одноконечных нарциссических профилей предпочтений. В нашей работе определяются групповые отделимые нарциссические профили предпочтений и подсчитываются.

Единственный известный результат относительно числа групповых сепарабельных профилей предпочтений – это неравенство Лакнера и Лакнера [79]: $\#^{GSPP}(n,m) \leq m!\left(3+2\sqrt{2}\right)^{m(n-1)}$.

Этот подраздел имеет два основных результата. Во-первых, получена точная формула для числа группоразделимых профилей предпочтений (Теорема 1.10). Комбинаторное доказательство основано на биекции между разделяемыми группами доменами и путями Шредера без пиков на уровне один, подсчитанных малыми числами Шредера. Во-вторых, получена рекурсивная формула для числа нарциссических группоразделимых профилей предпочтений (Теорема 1.11).

1.4.2 Модель и основные результаты

Пусть конечное множество $X = \{1, \dots, m\}$ будет множеством альтернатив и конечное множество $\mathcal{N} = \{1, \dots, n\}$ будет множеством избирателей. Каждый избиратель $i \in \mathcal{N}$ имеет линейный порядок P_i на X. Первый элемент является наилучшим, а последний – наихудшим. Пусть $\mathcal{L}(X)$ будет множеством всех линейных порядков на X. Кортеж из n порядков предпочтений является профилем предпочтений $\mathcal{P} = (P_1, \dots, P_n) \in \mathcal{L}(X)^n$. Функция $pos(P_i, j) = |\{x \in X | xP_ij\}| + 1$ отражает позицию альтернативы j в порядке P_i.

Профиль предпочтений $\mathcal{P}$ является отделимым если для каждого $A \subseteq X$, $|A| \geq 2$ существует разбиение на два непустых множества $A', A \backslash A'$ таких что для каждого $i \in \mathcal{N}$, мы имеем либо $aP_i b$ для каждого $a \in A'$, $b \in A \backslash A'$, либо $bP_i a$ для каждого $a \in A$, $b \in A \backslash A'$.

Домен $\mathcal{D} \in \mathcal{L}(X)^d$ представляет собой набор из d различных порядков предпочтений. Мы говорим, что профиль предпочтений принадлежит домену, если все заказы предпочтений из этого профиля предпочтений принадлежат домену.

Домен предпочтений является максимальным доменом Кондорсе, если каждый профиль предпочтений с нечетным числом избирателей из этого домена не дает циклов в простом соотношении большинства. Домены Кондорсе широко изучаются в теории общественного выбора (см. [81] для обзора этой области).

Домен предпочтений симметричен, если он содержит обратный порядок каждого порядка предпочтений в домене. Домен, который является максимальным доменом Кондорсе, а также симметричным доменом, является максимальным симметричным доменом Кондорсе (для подробного изучения этого свойства см. [82].

Домен предпочтений минимально богат, если для любой альтернативы $x \in X$ существует порядок $P \in \mathcal{D}$ такой, что P имеет x в качестве первой альтернативы. Минимальное богатство играет важную роль в характеристике однопиковых доменов предпочтений [83].

Домен предпочтений является нормальным, если он содержит порядок 123...m и противоположный ему [82].

Домен предпочтений можно разделить на группы, если все профили предпочтений из этого домена можно разделить на группы. Каждый домен скобок разделяется на группы. Каждый разделяемый группой профиль предпочтений принадлежит соответствующему домену скобок.

Скобки с одной скобкой генерируют домен с двумя обратными порядками. Любые дополнительные скобки удваивают количество предпочтительных заказов в домене. Домен скобок имеет 2^f порядков предпочтений, где f-количество скобок. Максимальный разделяемый группой домен является разделяемым группой доменом таким, что любой другой домен более высокой мощности, включающий этот домен, не является разделяемым группой доменом. Скобки, которые генерируют максимальный разделяемый группой домен, называются двоичными скобками. Таким образом, каждая максимальная группоразделимая область содержит 2^(m-1) порядков. Каждый разделяемый группой профиль предпочтений принадлежит максимальному разделяемому группой домену.

Число помеченных двоичных скобок слова с m буквами равно числу максимальных группоразделимых доменов. Из [84], число таких скобок и число максимальных группоразделимых доменов равно

$$\#^{MGSD}(m) = \prod_{i=2}^{m}(2i-3), \text{ где } \#^{MGSD}(1) = 1.$$

Теорема 1.10 [85]. Число отделимых профилей предпочтений равно

$$\#^{GSPP}(n,m) = m!\sum_{k=0}^{m-2}\frac{(-1)^k}{m-1}\binom{m-1}{k}\binom{2m-2-k}{m}2^{(m-k-1)(n-1)}.$$

Теорема 1.11 [85]. Число отделимых нарциссических профилей предпочтений удовлетворяет следующему условию:

$$\#^{\mathrm{NGSPP}}(\mathrm{n}) = \sum_{\mathrm{k}=1}^{\mathrm{n}-1}\binom{\mathrm{n}-1}{\mathrm{k}-1}2^{2\mathrm{nk}-2\mathrm{k}^2-\mathrm{n}}\#^{\mathrm{NGSPP}}(\mathrm{k})\#^{\mathrm{NGSPP}}(\mathrm{n}-\mathrm{k}),$$

где $\#^{\mathrm{NGSPP}}(1) = 1$.

1.4.3 Некоторые выводы

В данном подразделе показано, что группоразделимые домены являются подклассом максимальных симметричных доменов Кондорсе и что все максимальные симметричные домены Кондорсе наибольшей мощности являются

группоразделимыми. Все представленные примеры максимальной симметричной области Кондорсе имеют мощность 2^{k} для некоторых $k \in N$.

В этом подразделе соединяются групповые разделяемые ссылочные домены, упорядоченные деревья и пути Шредера, и это обеспечивает новые биекции. Эти биекции приводят к новому комбинаторному результату перечисления группоразделимых профилей предпочтений. Этот результат принципиально решает проблему, впервые сформулированную в [79].

1.5 Исследование условий согласованной единственности стабильного паросочетания

1.5.1 Введение

В данном подразделе изучается классическая задача о марьяже [86]. Экхоут [87] нашел достаточное условие единственности стабильного паросочетания. Следуя Кларку [88] это условие называется условием последовательных предпочтений (далее УПП). На текущий момент это наиболее часто применимое условие единственности в этой области.

Условие отсутствия пересечения [88], скоординированные предпочтения [89], условие универсальности [90], условия ацикличности [91, 92, 93] являются более строгими, чем УПП.

Существует несколько обобщений УПП для более общих задач о паросочетаниях. Итеративная α-сводимость [92], обощенное УПП [93], условие Ву [94] расширяют УПП для случая неприемлемых пар и/или множеств неравной мощности.

Обобщение УПП остается открытой проблемой. α-условие разработанное в [95] является первым существенным расширение УПП. В работе продемонстрировано, что профиль предпочтений удовлетворяет α-условию тогда и только тогда профиль имеет единственное стабильное паросочетание во всех подмножествах не нарушающих общего паросочетания.

1.5.2 Задача о марьяже

Рассмотрим два конечных множества агентов: множество $\mathcal{F}$ и множество $\mathcal{M}$, $\#\mathcal{F} = n, \#\mathcal{M} = m$. Типичный агент из $\mathcal{F}$ обозначен как A, B, C и типичный агент из $\mathcal{M}$ как a, b, c. Альтернатива остаться без пары обозначена как $\emptyset$. Мы рассматриваем две постановки задачи: с недопустимыми парами и без недопустимых пар.

Все агенты имеют предпочтения на множестве других агентов. Предпочтения являются полными и транзитивными. Мы рассматриваем только строгие предпочтения. Проблема построения стабильного паросочетания это тройка $(\mathcal{F}, \mathcal{M}, \succ)$, состоящая из двух множеств и предпочтений. Паросочетание это соответствие μ: $\mathcal{F} \cup \mathcal{M} \to \mathcal{F} \cup \mathcal{M} \cup \{\emptyset\}$, такое что $\forall x \in \mathcal{F} \cup \mathcal{M}$, если $\mu(x) \neq \emptyset$, тогда $\mu\big(\mu(x)\big) = x$. В примерах паросочетания представлены как список пар. Паросочетание является стабильным, если оно не заблокировано никакой парой $(X, x) \in \mathcal{F} \times \mathcal{M}$.

1.5.2.1 Парадокс разбиения рынка

Рассмотрим следующий профиль предпочтений

$$\begin{array}{ll} (b, a, c)_A; & (A, C, B)_a; \\ (a, b, c)_B; & (B, A, C)_b; \\ (a, c, b)_C; & (C, A, B)_c, \end{array}$$

который имеет единственное стабильное паросочетание $(a, A), (b, B), (c, C)$. После исключения пары (c, C) получим профиль предпочтений

$$\begin{array}{l} (b, a)_A; (A, B)_a; \\ (a, b)_B; (B, A)_b, \end{array}$$

который имеет два стабильных паросочетания. Мы называем этот феномен парадокс разбиения рынка. Даже в случае единственного стабильного паросочетания разбиение рынка ведет к множеству стабильных паросочетаний. Агент C в данном примере является интерраптором в терминах [96]. Следующее свойство защищает от возникновения парадокса.

Профиль предпочтений P удовлетворяет свойству согласованной единственности, если:

– профиль предпочтений имеет единственное стабильное паросочетание μ;

– для любого подмножества агентов ограничение профиля на это множество с соответствующими парами имеет единственное стабильное паросочетание.

Согласованная единственность является более строгим требованием, чем единственность стабильного паросочетания, но более слабым, чем единственность во всех подмножествах [88]. Согласованность является важным свойством паросочетаний (см. [97, 98]). Согласованная единственность определяет домен предпочтений, для которых стабильные правила являются согласованными.

1.5.2.2 α-условие

Назовем пару агентов, для которых партнер является наилучшим, фиксированной парой [88]. Экхоут [87] рассматривал модель без запрещенных пар и определил следующее условие.

Условие последовательных предпочтений (УПП) [87]. Существует порядок $\mathcal{F}, \mathcal{M}$, такой что

$$\forall X_i \in \mathcal{F}: \forall j > i,\ x_i \succ_{X_i} x_j,$$

$$\forall x_i \in \mathcal{M}: \forall j > i,\ X_i \succ_{x_i} X_j.$$

УПП является достаточным условием единственности стабильного паросочетания.

α-условие. Существует стабильное паросочетание μ и возможно упорядочить агентов из $\mathcal{F}$ так, что

$$\forall X_i \in \mathcal{F}: \forall j > i, \qquad \mu(x_i) \succ_{X_i} x_j,$$

и возможно упорядочить агентов из $\mathcal{M}$ так, что

$$\forall x_i \in \mathcal{M}: \forall j > i,\ \mu(X_i) \succ_{x_i} X_j.$$

Следующий пример проясняет разницу между УПП и α-условием.

Пример 1.5. Следующий профиль предпочтений

$$\begin{array}{ll} (a,c,b)_A, & (C,A,B)_a, \\ (a,b,c)_B, & (B,A,C)_b, \\ (b,c,a)_C, & (B,C,A)_c. \end{array}$$

имеет единственное стабильное паросочетание $(a,A),(b,B),(c,C)$, но не удовлетворяет УПП. Имея порядок $x_1 = a, X_1 = A,\ x_2 = b, X_2 = B,\ x_3 = c, X_3 = C$, и второй порядок $x_1 = b, X_1 = B,\ x_2 = C, X_2 = C,\ x_3 = a,\ X_3 = A$ мы показываем, что профиль удовлетворяет α-условию.

Теорема 1.12 [95]. α-условие является достаточным условием согласованной единственности.

Теорема 1.13 [95]. При отсутствии недопустимых пар α-условие является необходимым и достаточным условием согласованной единственности.

1.5.3 Некоторые выводы

На рисунке 1.5 представлены доли профилей предпочтений, для которых выполняются УПП, α-условие и единственность стабильных паросочетаний. Результаты получены путем генерации 1 000 000 независимых случайных профилей предпочтений. Из рисунка 1.5 мы видим, что α-условие все еще довольно ограничительно, но оно значительно слабее, чем УПП. При $n \geq 15$ число профилей предпочтений, для которых выполняется α-условие, более чем в 10 раз больше, чем для УПП. Эти условия имеют различные порядки сходимости.

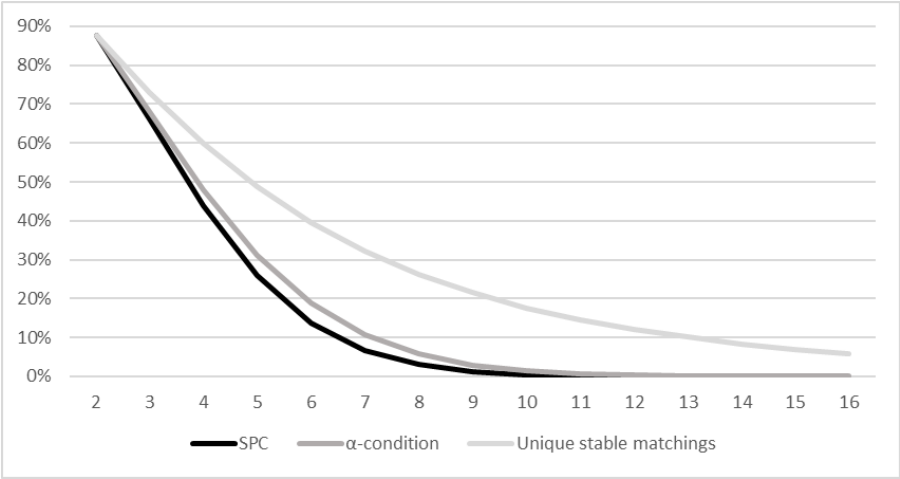


Рисунок 1.5 – Случай $n = m$ и отсутствие недопустимых пар

2 Разработка, исследование и применение методов анализа данных

В этом разделе отчета приведены результаты по разработке, исследованию и применению методов анализа данных. Полученные результаты отражены в следующих подразделах:

– исследования по сетевому анализу данных;

– исследования по анализу текстов и кластерному анализу;

– разработка методов анализа паттернов.

2.1 Исследования по сетевому анализу данных

2.1.1 Введение

В данной части исследовательского проекта затрагиваются вопросы о методах оценки влияния в сложных сетевых структурах, выполнении некоторых свойств и аксиом для классических мер центральностей, а также выявлении ключевых элементов в прикладных задачах сетевого анализа.

Большинство существующих систем (социальные, технологические, биологические и т.д.) имеет сложную сетевую структуру, одни и те же вершины в которой могут иметь различные виды взаимосвязей друг с другом. Такой тип структур называется мультиплексным, который можно определить через набор из L графов $\{G_l = (V, E_l)\}_{l=1}^{L}$, где $E_l \subseteq V \times V$ – набор ребер, или как тройку $G_M = (V, E, L)$, где V – набор вершин, L – набор слоев, описывающих разные типы взаимоотношений, $E \subseteq V \times V \times L$ – набор взвешенных ребер [99, 100, 101, 102]. Дополнительно, каждый слой сетевой структуры может быть описан с помощью взвешенной матрицы связности $W^l = [w_{ij}^l]$, а каждая вершина может характеризоваться набора параметров $Q_i = (q_i^1, q_i^2, \dots, q_i^L)$, которые могут показывать порог влияния на данную вершину на каждом слое.

Одним из важных аспектов сетевого анализа является вопрос определения её наиболее влиятельных и уязвимых элементов. Так как большинство существующих сетей имеют более сложную структуру, вопрос определения

наиболее важных и уязвимых элементов гораздо сложнее, чем для сетей, имеющих один тип связи между вершинами. На данный момент существует ряд моделей, возникших на основе классических индексов центральности, которые позволяют оценить влияния в многослойных сетевых структурах. Одним из таких подходов является оценка центральности вершин на каждом индивидуальном слое, которая затем агрегируется с учетом веса вершин или других процедур агрегирования. Тем не менее, данный подход не учитывает информации о том, как две вершины связаны между собой на различных слоях. Второй подход заключается в преобразовании многослойной сетевой структуры в однослойную, в которой вес на ребрах между вершинами вычисляется как линейная свертка информации на ребрах каждого слоя [99, 103]. К сожалению, данный метод не учитывает информацию о структуре сети на каждом отдельном слое. Более того, взаимосвязи между вершинами могут быть даны в разных измерениях, что снижает интерпретируемость агрегированного веса ребра. Существует и более сложные подходы, которые заключаются в построении матрицы обобщенной инцидентности, где различные слои сетевой структуры связываются между собой, а влияние определяется на основе вычисления собственного вектора построенной матрицы [104, 105, 106, 107, 108]. Основная идея данного подхода заключается в том, что центральность вершины зависит не только от взаимосвязей на каком-то одном слое, но и от взаимосвязей на всех других слоях. В [109] было показано, как некоторые существующие индексы центральности могут быть адаптированы к сетям данного типа. Тем не менее, во многих приложениях не вполне понятно, как связать между собой различные слои между собой, так как влияние на одном слое сети не всегда приводит к влиянию на других слоях сетевой структуры. В связи с этим, вопрос построения матрицы обобщенной инцидентности является спорным и требует более детального изучения того, как различные слои связаны друг с другом.

Следует также отметить, что большинство моделей по оценке важности и влияния в сетях анализируют только узловые элементы и не оценивают важность связей между ними. При этом, выявление наиболее сильных и важных связей также помогло бы нам лучше понять устройство нашей сетевой структуры.

Существует несколько методов по оценке рёберной важности в сетях. Наиболее известная мера – рёберная промежуточность (edge betweenness), которая, по аналогии с вершинной промежуточностью, выделяет те рёбра, через которые проходит больше всего кратчайших путей [110]. Но такая мера даёт значимый результат в тех сетях, в которых важно учитывать длины кратчайших путей, таких как транспортных сетей. Другой подход по оценке важности рёбер описан в [111] и называется k-path рёберная центральность. Данная мера основана на случайных блужданиях в графах, где берутся пути длиной не более чем k. Данный подход может быть применён к сетевым моделям распространения. В [112] представлен подход, основанный на функциях надёжности: ребро e_1 лучше (в некотором смысле), чем ребро e_2, если функция надёжности меньше без ребра e_1, чем без ребра e_2. Также важность рёбер может быть выявлена на основе потоковых моделей в сетях, таких как поиск максимального потока, минимального разреза, пропускной способности ребра и т.д. [113]. Подход оценки важности рёбер, основанный на построении остовных деревьев описан в [114]. Более того, классические меры центральности также могут быть применены к графу для оценки важности рёбер, если исходный граф представить в виде линейного графа, т.е. графа, в котором рёбра играют роль узлов, а узлы – роль рёбер. Однако, такой подход имеет высокую сложность.

В данной работе важность рёбер оценивается на основе непрямого влияния между вершинами. Такой подход помогает выявить неявные каналы влияния в сетевых структурах, которые обеспечивают проводимость более высокого влияния между вершинами.

Ещё одна часть исследования посвящена анализу выполнимости аксиом для мер центральности в сетях. Эта работа является продолжением, описанным в пункте 4.1.9 [1].

Сетевой анализ используется для изучения распространения информации, эпидемий в обществе, устойчивости финансовых потоков [115, 116]. Кроме того, анализируется распределение влияния в сети для выделения наиболее важных групп в выборных органах [117]. Террористические сети исследуются для выделения связей между группировками в [118, 119], транспортные сети, их эффективность и распределение нагрузки анализируются в [120, 121], существуют так же и другие приложения. Одна из основных задач в таких исследованиях — выделение наиболее влиятельных вершин или их объединений.

Выбор решения для этой задачи зависит от постановки. Например, для определения людей, включенных в рекламную кампанию некоторого товара с целью наибольшего распространения информации о нем, нужна одна мера, а для определения наиболее релевантной странице в Интернете при поиске информации, подойдет другая. Исследование свойств различных мер влияния для сетей помогает выбрать наиболее подходящее решение в каждом случае.

Мы исследуем два основных подхода к определению важных узлов сети. Во-первых, меры центральности для сетей: как классические меры для определения центральности, так и более сложные, основанные на поиске собственных векторов матриц смежности графов. Во-вторых, мы рассматриваем теоретико-игровой подход к распределению общего выигрыша между участниками в кооперационных играх, на основе которого так же определены меры влияния вершин в сетях [122]. В следующих разделах описаны свойства и аксиомы для мер центральности, основанных на структуре сети, случайном блуждании и симуляциях, на решениях для коалиционных игр при ограниченной кооперации.

Другая часть исследования посвящена прикладным задачам сетевого анализа, в частности задаче моделирования международных миграционных процессов, а также вопросу международной торговли.

Международная миграция является важнейшей проблемой современного общества. Численность международных мигрантов в мире растет ежегодно и в настоящий момент составляет 272 миллионов человек [123]. Число международных мигрантов с 2010г. выросло на 51 миллион человек, при этом число беженцев увеличилось на 13 миллионов с 2010 г. и составило около четверти от роста числа международных мигрантов [123]. В современном мире появляется необходимость в новых подходах к моделированию международной миграции, т.к. важность этого процесса возрастает, как и число стран, вовлеченных в международную миграцию.

Моделирование международной миграции возможно с помощью сетевого подхода [124, 125]. В данной модели вершинами графа являются страны, а ребрами – потоки международных мигрантов. Данная работа продолжает исследование лаборатории [1] и расширяет базу данных за счет новых потоков 2015-2017 гг. Основной частью исследования этого года является дополнение базы данных новыми периодами.

Что касается вопроса международной торговли, в последние десятилетия сетевой анализ этой области получил широкое распространение прежде всего благодаря расширению доступа к базам данных по международной торговле, таким как база статистических данных Организации Объединенных Наций по торговле товарами (англ. the United Nations International Trade Statistics Database), база данных Международного торгового центра (англ. the International Trade Center Database), база данных Европейского статистического бюро по международной торговле товарами и услугами (англ. European Statistical Office International Trade in Goods and Services Database) [126]. Однако, несмотря на распространение сетевого анализа в международной торговле, в большинстве исследований анализируются только прямые взаимодействия между странами на различных базах данных.

Данная исследовательская работа анализирует прямые и непрямые взаимодействия в международной торговой сети товарами в рамках нескольких задач:

1) обзора научной литературы по сетевому анализу международной торговли товарами;
2) описания классических индексов центральности и новых индексов центральности;
3) сбора данных о международной торговле товарами;
4) обработки собранных о международной торговле товарами данных;
5) расчета классических индексов центральности и новых индексов центральности;
6) сравнения расчетов классических индексов центральности и новых индексов центральности.

2.1.2 Оценка влияния в мультиплексных сетях

2.1.2.1 Описание модели

В нашей работе предлагается модель оценки влияния в мультиплексных сетях с использованием правил коллективного выбора. Построенная модель по большей части разработана для ситуаций, когда влияние на различных слоях между вершинами происходит независимо друг от друга или когда нет четкой информации о взаимосвязи между слоями. В таком случае задача оценки влияния в многослойных сетях может быть сведена к стандартной задаче коллективного выбора. Основная идея нашего подхода заключается в перестроении исходной сети в сеть влияния, определении попарного влияния вершин друг на друга и агрегирования данной информации с использованием концепций теории коллективного выбора.

Преобразование исходной сети и оценка попарного влияния основано на ранее разработанных моделей влияния (индексы дальних взаимодействия LRIC), предложенных в работах [127, 128, 129], которые позволяют учитывать

индивидуальные атрибуты вершин, возможности их группового влияния, а также непрямые взаимодействия между собой.

Процедура оценки попарного влияния делится на два этапа. На первом этапе исследуется влияние на каждую индивидуальную вершину отдельно, в рамках чего прямое влияние одной вершины на другую заключается в поиске минимальной группы, совокупная интенсивность которой превосходит ранее определенный порог влияния. Значение прямого влияния между вершинами измеряется от 0 (отсутствие влияния) до 1 (максимальное влияние). Тем не менее, так как вершины могут влиять друг на друга не только напрямую, но через какие-то промежуточные узлы, второй этап модели заключается в рассмотрении различных каналов влияния вершин друг на друга. В работах [127, 128, 129] было предложено несколько моделей оценки непрямого влияния, например, путем рассмотрения всех возможных путей между вершинами или нахождении пути, имеющем наибольшую силу влияния. Таким образом, предложенная модель преобразует исходную сеть в сеть попарного влияния вершин друг на друга на каждом отдельном слое. Обозначим через C_M полученный набор матриц, характеризующих попарное влияние вершин на конкретном слое.

Для агрегирования информации влияния вершин друг на друга на различных слоях мы будем использовать ординальный подход, который не чувствителен к проблеме несравнимости различных слоев. Так как мы рассматриваем взаимодействия между вершинами на каждом слое независимо, т.е. не рассматривается ситуация, когда влияние на вершину на одном слое приводит к влиянию на другом слое, можно преобразовать исходную задачу оценки влияния в мультиплексных сетях к задаче коллективного и многокритериального принятия решений. Основной целью данной задачи является оценка того, какой набор альтернатив необходимо выбрать из исходного набора, характеризуемого некоторым набором критериев. Так как попарное влияние дано в количественной шкале, мы можем рассмотреть данные значения как некоторые значения функции полезности и связать задачу влияния в сетевых структурах

с задачей коллективного выбора на основе этих функций полезности. Для решения данной проблемы мы адаптируем ординальный подход, основанный на мажоритарном отношении и построении турнирных матриц, который является более устойчивым к проблемам противоречия критериев друг с другом, а также их пороговой некомпенсируемости.

Определение 2.1. Определим общее число слоев, в котором вершина i имеет большее влияние на вершину j по следующей формуле

$$n(i,j,C_M) = card\{l \in N \,|c_{ij}^l - c_{ji}^l > 0\}, \tag{2.1}$$

Отметим, что данная формула может принимать более универсальный вид путем задания некоторого параметра ε>0

$$n(i,j,C_M) = card\{l \in N \,|c_{ij}^l - c_{ji}^l > \varepsilon\}. \tag{2.2}$$

Определение 2.2. Определим мажоритарное отношение μ по следующей формуле

$$i\mu j \Leftrightarrow n(i,j,C_M) > n(j,i,C_M). \tag{2.3}$$

Другими словами, будем считать, что вершина i доминирует вершину j, если общее число слоев, где вершина i оказывает большее влияние на вершину j, больше общего числа слоев, где наблюдается обратная ситуация.

Существует множество способов того, как произвести ранжирование вершин на основе формул (2.1)-(2.3). В Таблице 2.1 представлен список некоторых процедур коллективного выбора.

Таблица 2.1 – Список процедур коллективного выбора

#	**Название процедуры выбора**	**Тип процедуры**
1	Правило Борда	Позиционные правила
2	Правило Блэка	
3	Обратное правило Борда	
4	Правило Нансона	

Продолжение таблицы 2.1

#	**Название процедуры выбора**	**Тип процедуры**
5	Победитель Кондорсе	Правила, использующие мажоритарное отношение
6	Минимальное доминирующее множество	
7	Минимальное непокрытое множество	
8	Минимальное слабоустойчивое множество	
9	Правило Фишберна	
10	Непокрытое множество I	
11	Непокрытое множество II	
12	Правило Ричалсона	
13	Ядро	
14	k-устойчивые множества	
15	Первое правило Копланда	Правила, использующие вспомогательную числовую шкалу
16	Второе правило Копланда	
17	Третье правило Копланда	
18	Минимаксная процедура	Правила, использующие турнирную матрицу
19	Процедура Симпсона	

Отметим, что для оценки наиболее важной вершину могут использоваться и другие модели коллективного выбора. В работах [130, 131] приведено описание различных правил коллективного выбора, а также исследуются нормативные свойства данных процедур. Матрично-векторное представление различных процедур, использующих мажоритарное отношение μ приведено в [132].

2.1.2.2 Применение модели к сети торговли продовольствием

Рассмотрим применение предложенной модели оценки влияния в мультиплексной сети, характеризующую информацию о международной торговле по наиболее важным для продовольственной безопасности продуктам питания: мясо птицы, пшеница, рис.

Вершинами сети являются страны, а взаимосвязи показывают общий объем торговли между странами по отдельному продукту. Вершины сети неоднородны, так как они имеют свой набор характеристик (например, ВВП, население, уровень потребления и производства и т.д.). Другими словами, один и тот же уровень торговли может быть критическим для одной страны и незначительным для другой страны. В связи с этим необходимо учитывать ин-

дивидуальные атрибуты стран для оценки общего влияния в сети международной торговли. Для оценки попарного влияния с учетом данных особенностей, мы применили индекс дальних взаимодействий LRIC по методологии, описанной в работах [133, 134]. Как результат, была рассчитана информация о попарном влиянии вершин друг на друга на каждом рассматриваемом слое и агрегируем данную информацию в соответствии с предложенной моделью (см. Таблица 2.2).

Таблица 2.2 – Ранжирование стран (ТОП-10)

#	Правило Борда	Первое правило Копленда	Минимальное слабоустойчивое множество
1	США	США	США
2	Тайланд	Тайланд	Тайланд
3	Бразилия	Бразилия	Бразилия
4	Аргентина	Аргентина	Аргентина
5	Канада	Канада	Канада
6	Уругвай	Уругвай	Уругвай
7	Индия	Индия	Индия
8	Вьетнам	Пакистан	Пакистан
9	Италия	Вьетнам	Вьетнам
10	Франция	Австралия	Австралия

Как показано в Таблице 2.2, наиболее важным игроком на рынке торговле рассматриваемых 3 продуктов питания является США. Действительно, данная страна является основным экспортером по всем рассматриваемым продуктам питания. Индия, Пакистан, Тайланд и Вьетнам являются основными экспортерами риса, что объясняет их лидирующие позиции (Тайланд также входит в ТОП-15 по экспорту мяса птицы). Что касается других стран, Аргентина, Бразилия и Канада имеют большие значения экспорта мяса птицы и пшеницы. Наконец, можно обратить внимание, что данные ранжирования коррелируют друг с другом и показывают схожие результаты и позволяют выявить наиболее важных участников торговли с учетом рассматриваемых типов продуктов.

2.1.2.3 Применение модели к сети торговли вооружением

Рассмотрим также применение предложенной модели оценки влияния к сети торговли вооружением. В качестве исходных данных была взята база данных SIPRI Arms Transfers Database, которая была предоставлена Стокгольмским институтом исследования проблем мира [135]. База данных содержит информацию по всем операциям за период с 1950 по 2018 годы с указанием типа поставленного вооружения и его стоимости (оценка производилась Стокгольмским институтом, TIV). Таким образом, вершинами сети являются страны, а взаимосвязи показывают объем торговли между странами по отдельному типу вооружения. Торговля вооружения поделена на следующие виды: системы противовоздушной обороны, самолеты, бронетехника, артиллерия, двигатели, ракеты, военно-морское оружие, спутники, датчики, корабли и другое оружие. Информация о торговле вооружением по типу оружия приведена на рисунке 2.1.

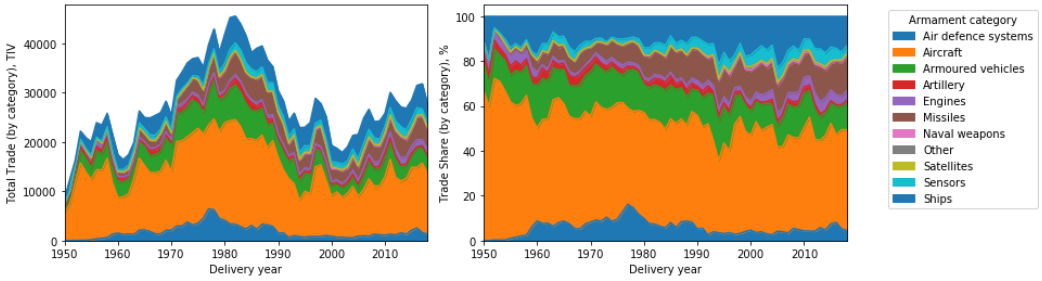


Рисунок 2.1 – Структура торговли вооружением по типу вооружения (объем и структура)

Как и в случае с торговлей продовольствия, вершины сети также неоднородны, так как они имеют свой набор характеристик. В связи с этим для оценки попарного влияния, мы применили также индекс дальних взаимодействий LRIC. Как результат, была рассчитана информация о попарном влиянии вершин друг на друга на каждом рассматриваемом слое за отдельный период и агрегируем данную информацию в соответствии с предложенной моделью. Полученные результаты, где итоговое агрегирование осуществляется по правилу Борда и первому правилу Коупленда, показаны на рисунке 2.2.

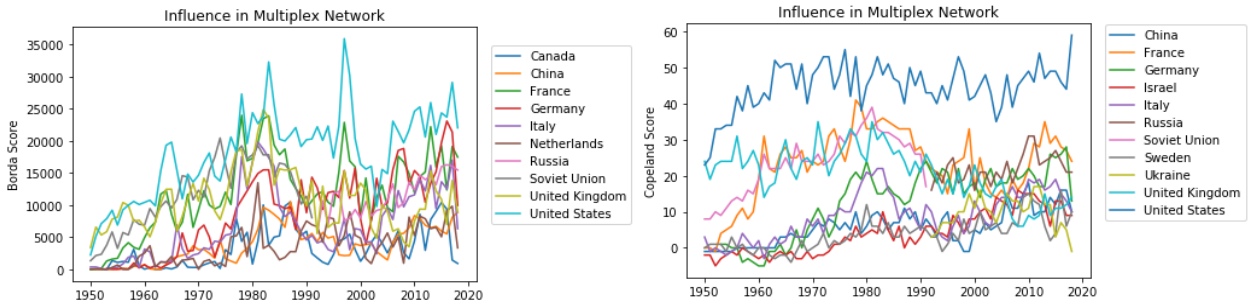


Рисунок 2.2 – Влияние в многослойной сети торговли вооружением

Как показано на рисунке выше, США, Великобритания, СССР (позже – Россия) и Франция являются наиболее влиятельными участниками в сети торговли вооружением. Например, США за различные периоды доминируют 30-50 стран по мажоритарному отношению, которое было построено на основе того, как страны влияют друг на друга на различных уровнях. Мы также можем увидеть, что относительное влияние Великобритании становится все ниже с течением времени. Тем не менее, можно также увидеть, что некоторые страны (Канада, Китай, Италия и т.д.) входят в список наиболее влиятельных стран за отдельные периоды. Наконец, можно обратить внимание, что ранжирования коррелируют друг с другом и позволяют выявить наиболее важных участников торговли вооружением.

2.1.3 Меры влияния связей в сетевых структурах

В данной части исследовательского отчёта описаны меры по оценки важности рёбер в сетях. Мы анализируем взвешенные направленные графы, где веса рёбер принимают значения от 0 до 1. Данной нормировки можно добиться, например, с помощью описанного выше подхода преобразования исходной сети в сеть интенсивностей прямого влияния.

Для описания модели сначала рассмотрим взвешенный направленный граф на трёх вершинах (см. рисунок 2.3).

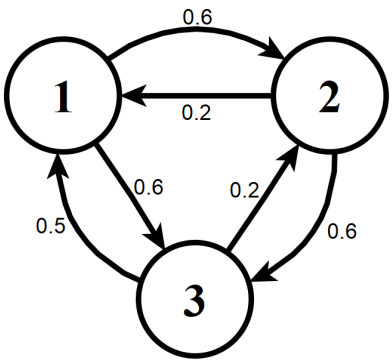


Рисунок 2.3 – Пример графа с интенсивностями прямого влияния

Предположим, что веса рёбер обозначают интенсивность прямого влияния. Наша задача определить ребро (или множество рёбер), через которые проходят наибольшие интенсивности непрямого влияния. Для этого мы рассмотрели меры из [127] по оценке непрямого влияния через всевозможные простые пути в сети. Один из подходов по оценке непрямого влияния через конкретный путь основан на перемножении интенсивностей прямого влияния на рассматриваемом пути.

Для графа на рисунке 2.3 мы можем посчитать прямое и непрямое влияние между всеми парами вершин. Так, вершина 2 влияет напрямую на вершину 1 с интенсивностью 0.2, а через вершину 3 (т.е. через путь 2–3–1) с интенсивностью $0.6 \cdot 0.5 = 0.3$, что больше, чем прямое влияние. Аналогично, вершина 3 влияет на вершину 2 напрямую с интенсивностью 0.2, а через путь 3–1–2 – с интенсивностью $0.6 \cdot 0.5 = 0.3$, что также больше, чем прямое влияние. Влияние остальных пар вершин вдруг на друга напрямую больше, чем через посредников (рисунок 2.4.). Тем не менее, можно заметить, что ребро (3,1) два раза встречалось на непрямых путях, которые увеличивали общее влияние вершин друг на друга по сравнению с прямым влияние. Следовательно, можно сказать, что ребро (3,1) играет важную роль в проводимости непрямого влияния в сети. Примечательно, что классическая мера рёберной промежуточности (edge-betweenness) всем рёбрам даёт одинаковый вес. Другие известные подходы также не выделяют это ребро как наиболее важное.

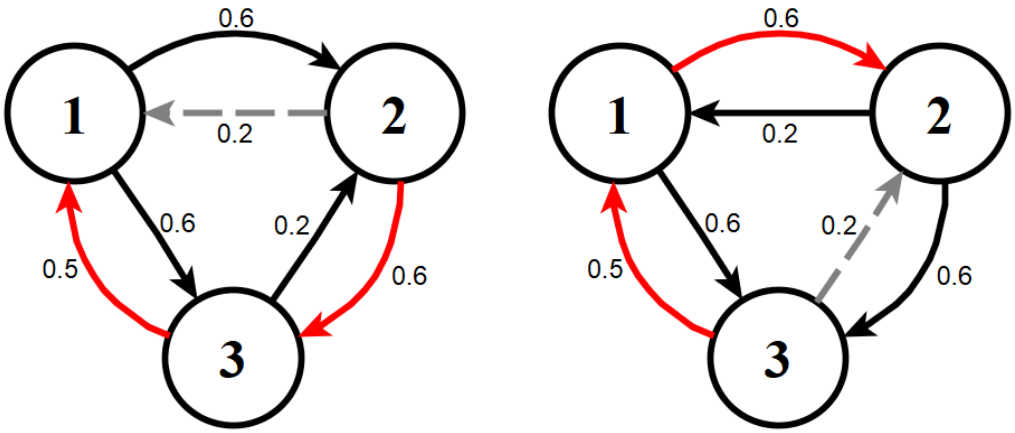


Рисунок 2.4 – Влиятельные непрямые пути в графе

Далее, опишем модель по оценке влияния рёбер в сети формально. Имеется граф $G = (V, E)$, V – множество вершин $\{1, \dots, n\}$, $E = \{e_{ij}\}$ – множество взвешенных направленных рёбер. Обозначим за P_{ij}^k -ый путь между вершинами i и j длиной $|P_{ij}^k| = l$: $P_{ij}^k = \left(e_{it_1}, e_{t_1t_2}, \dots, e_{t_{l-1}j}\right)$, где $t_1, \dots, t_{l-1}$ – вершины на этом пути. Для каждого пути можно посчитать функцию непрямого влияния через этот путь $Inf(P_{ij}^k)$, например, как произведение прямых влияний вдоль этого пути (как в примере выше)

$$Inf(P_{ij}^k) = e_{it_1} \times e_{t_1t_2} \times \dots \times e_{t_{l-1}j} \tag{2.4}$$

или как минимальное значение прямого влияния вдоль этого пути, что соответствует пропускной способности общего влияния через этот путь:

$$Inf(P_{ij}^k) = \min\left(e_{it_1}, e_{t_1t_2}, \dots, e_{t_{l-1}j}\right). \tag{2.5}$$

Теперь наша задача может быть сформулирована следующим образом: для каждой пары вершин необходимо выявить тот непрямой путь, влияние через который будет больше, чем любое другое влияние, в том числе и прямое влияние вершин друг на друга:

$$k_{ij}^* = \arg\max_k Inf\left(P_{ij}^k\right) \qquad w.r.t.\ Inf\left(P_{ij}^k\right) > e_{ij}. \tag{2.6}$$

Если мы рассматриваем пути определённой длины (например, все пути, длиной не более, чем l), то данная задача может быть решена аналогично задаче возведения матрицы смежности A в степень l, т.к. A^l хранит число путей длины l между вершинами. Более конкретно, когда мы умножаем матрицу A на матрицу B, мы последовательно умножаем элементы в соответствующей строке матрицы A на соответствующий столбец в матрице B, после чего суммируем произведения. Если мы хотим выявить наиболее влиятельные пути по формулам (2.4) и (2.6) мы также можем применить данный подход, заменяя суммирование произведением взятием максимума из произведений. А если мы хотим выявить влиятельные пути по формулам (2.5) и (2.6), то мы произведение элементов заменяем функцию минимума, а суммирование – на функцию максимума.

Обозначим множество решений задачи (2.6) как $P^* = \left[P_{ij}^{k^*}\right]$. Данное множество содержит все пути, которые были выбраны как наиболее влиятельные. После того, как мы получили это множество решений задачи (2.6), мы можем оценить вклад каждого ребра, которые встречаются на выбранных путях.

Предлагается несколько подходов по тому, как оценивать значимость того или иного ребра. Некоторые из подходов описаны в Таблице 2.3.

Таблица 2.3 – Алгоритм оценки влияния рёбер в графе G

Algorithm EdgeScoring

INPUT: $G,\ P^*, P^{**}$ #G – граф
P^* - множество наиболее влиятельных путей
#P^{**} - множество вторых наиболее влиятельных путей

for e **in** $edges(G)$: #инициализация
$Score1[e] = 0$
$ScoreInf[e] = 0$
$ScoreContr[e] = 0$
$ScoreDiff[e] = 0$

for i, j **in** $nodes(G)$ # для каждой пары вершин
for e **in** P_{ij}^*: # для каждого ребра на выбранном пути между i и j

Продолжение таблицы 2.3

$Score1[e]$ + = 1 # +1
$ScoreInf[e]$ + = $Inf(P_{ij}^{*})$ # текущее значение влияния
$ScoreContr[e]$ + = $Inf(P_{ij}^{*})$ - $weight(i,j)$ #вклад непрямого пути
$ScoreDiff[e]$ + = $Inf(P_{ij}^{*}) - Inf(P_{ij}^{**})$ #разница между первым и
#вторым наиболее
#влиятельными путями

return $Score1, ScoreInf, ScoreContr, ScoreDiff$

Так, например, в переменной $Score1$ для каждой вершины хранится число раз, сколько вершина встретилась во множестве P^{*}. Данная мера показывает, как часто более влиятельные каналы проходили через каждую вершину. Другая переменная $ScoreInf$ вычисляется как сумма влияний через пути, в которые входит рассматриваемая вершина. Данная мера помогает оценить вклад каждой вершины в общее влияние через непрямые пути. Переменная $ScoreContr$ рассчитывается как сумма разностей между влиянием через непрямой путь, содержащим рассматриваемое ребро, и соответствующего прямого влияния. $ScoreDiff$ оценивает разность между первым по влиятельности непрямым путём, содержащих рассматриваемое ребро, и вторым по влиятельности непрямым путём между парами вершин. Данная мера выражает возможные потери в мере влияния при возможном удалении наиболее влиятельного пути. Все рассмотренные меры могут быть нормированы с учётом индивидуальных атрибутов рёбер (таких, как вес ребра или другие внешние атрибуты). Возможны и другие меры по учёту важности рёбер из множества P^{*}.

В результате, мы описали подход для определения наиболее влиятельных непрямых путей в графе, а также предложили несколько мер для оценки важности рёбер на данных путях. На графах специального вида данные меры дают ожидаемый результат. Например, в полном графе с одинаковыми весами рёбер или в графе типа звезды с направленными рёбрами алгоритм не выделяет никакое ребро как наиболее влиятельное. С другой стороны, алгоритм

позволяет выделить мосты при их наличии в графах. Алгоритм также может быть применён к многокомпонентным графам.

2.1.4 Критические группы в сетях и меры центральности: проверка выполнимости аксиом.

К классическим мерам центральности относятся степенная (degree), близости (closeness), по посредничеству (betweenness). Кроме того, существуют меры, основанные на матрице смежности для сети, которые позволяют учитывать при расчете центральности для некоторого узла показатель для его соседей. К ним относятся центральность по собственному вектору (eigenvector) [136], центральность Боначича [136], престиж Катца. Для направленных графов используются PageRank, HITS (Hubs and authorities — рейтинг агрегаторов и авторитетов).

Коалиционная (или кооперативная) игра — это пара (N, w) множества игроков N и характеристической функции $w: 2^N \to R, w(\emptyset) = 0$. Функция w приписывает некоторый выигрыш каждому подмножеству игроков (коалиции) $S \subseteq N, |S| = s$. Коалиция, состоящая из всех игроков, называется гранд-коалицией и значение функции $w(N)$ называется стоимостью игры. Решением такой игры является распределение этой стоимости между всеми игроками на основе их ценности для коалиций. Такое решение приписывает больший выигрыш более важным игрокам, следовательно, может быть использовано для определения важных игроков.

Классическим решением кооперативных игр является мера Шепли, предложенная в [137]. В работе [138] предложено обобщение для игр с ограниченной кооперацией, такие игры названы коммуникационными. В них ограничение на кооперацию задается ненаправленным графом. В [139] определена мера Шепли для игр, в которых ограничение на кооперацию задается направленным графом. Таким образом, мера Шепли может быть обобщена для поиска наиболее важных вершин в сети.

На основе решения Шепли для коалиционных игр разработаны несколько мер центральности для сетей. В [140] предложена Attachment centrality, разработанная на основе меры Майерсона. В [127] описаны меры ближнего и дальнего взаимодействий. Эти меры основаны на индексе влияния, учитывающем предпочтения агентов по созданию коалиций.

Выборный орган (например, парламент) может быть описан кооперативной игрой определенного типа, в которой характеристическая функция принимает только значения 0 и 1. Тогда коалиция будет выигрывающей, если ее выигрыш равен 1, и грандкоалиция всегда будет выигрывающей. Игрок называется ключевым в коалиции, если без него эта коалиция перестает быть выигрывающей (то есть получает выигрыш 0). Для таких игр разработаны индексы влияния, отражающие важность агента при принятии решений.

К классическим индексам влияния относят индекс Банцафа [141], рассчитывающийся в зависимости от числа коалиций, в которых игрок является ключевым, и индекс Шепли-Шубика [137], являющийся упрощенным вариантом меры Шепли (в соответствии с упрощенной характеристической функцией). Обобщенный индекс, учитывающий предпочтения агентов по созданию коалиций представлен в [142].

В [143] предложены три аксиомы, основанные на поведении мер центральности в некоторых графах определенных конструкций: аксиомы Size, Density, Score monotonicity (объема, плотности, монотонности). Выполнение этих аксиом проверено для 13 различных мер в сетях, показан обобщенный результат. Все три аксиомы выполняются только для степенной центральности. [144] сформулировали аксиомы, также основанные на значениях индексов для графов определенной конструкции (центр звезды, центральные вершины цепи, крайние узлы цепи, изолированный узел). Выполнимость аксиом проверена для некоторых классических мер.

Аксиоматика — это набор аксиом, полностью описывающий некоторую меру, и набор соответствующих теорем и доказательств того, что мера удовле-

творяет представленным аксиомам, и что это единственная мера, удовлетворяющая данному набору аксиом. В [139] представлена аксиоматика для меры Шепли в ориентированных графах. Для классических мер, таких как меры Шепли и Майерсона, существует несколько вариантов такого набора аксиом, см. [145, 138, 146, 147]. В [140] показана аксиоматика для разработанных мер центральности и степенной центральности, доказаны соответствующие теоремы.

2.1.5 Сетевой анализ международной миграции

Для анализа международной миграции между странами использовались данные по международным потокам мигрантов по странам мира. Данные по международным миграционным потокам были взяты из трех открытых источников [148, 149, 150]. Различные источники данных по международной миграции содержат различные списки стран, что позволяет максимально расширить число стран в анализе, используя при этом исключительно открытые источники данных.

Текущий анализ международной миграции позволил увеличить наблюдаемый период на три года с помощью баз данных Eurostat и OECD [149, 150]. База данных Eurostat содержит данные по потокам миграции в страны Европейского Союза, а также из данных стран в страны мира. Соответственно, база данных OECD основана по тому же принципу, для стран OECD. Период международной миграции 2015-2017, ставший доступным в недавнее время, имеет большое значение для анализа миграции между странами мира, т.к. охватывает миграционный кризис в Европейском Союзе.

Данные Eurostat по потокам эмиграции и иммиграции основаны на критерии гражданства мигранта. Данные 2015-2017 гг. состоят из потоков миграции между странами предоставившими статистические данные и странами мира. Данные OECD представлены 4576 потоками международной миграции по странам мира в 2017г. (Таблица 2.4).

Таблица 2.4 – Описание данных по миграции

База данных	Число потоков		
	2015	2016	2017
OECD	4764	4290	4576
Eurostat	5108	4987	5054

Помимо данных по потокам международной миграции также в работе использовались данные по ВВП (по паритету покупательной способности), населению, территории, доли агропромышленной составляющей в ВВП по 237 странам мира за 1970-2017гг. Данные по странам будут использованы для проведения исследования по сетевому анализу международной миграции.

Сетевой анализ в данной работе применяется с целью определения наиболее влиятельных стран в сети международной миграции в каждом временном промежутке. Индексы центральности (как классические, так новые, учитывающие непрямое взаимодействие между странами, групповое влияние и население принимающей страны [127]) позволяют выявить важность стран в сети миграции в различном контексте. В качестве иллюстрации использования индексов центральности в анализе международной миграции приведено ранжирование стран (Топ 10 по каждому индексу) за 2008 г.

В Таблице 2.5 приведены результаты анализа за 2008 г. Результаты классических индексов показывают, что наиболее привлекательными странами для мигрантов были преимущественно США, страны Евросоюза и Австралия (WInDeg). По оттоку мигрантов лидировали Германия, Китай и Великобритания (WOutDeg). Наибольшим влиянием в сети миграции с учетом влияния соседних стран обладали Китай, Германия и Польша (Eigenvector). Индексы, учитывающие ближние и дальние взаимодействия также выделили Китай и Польше и Германию.

Таблица 2.5 – Рейтинг влияния стран на основе индексов центральности (2008)

Country	WInDed	WOutDeg	Eigenvector	LRIC (Max)	SRIC
United States	1	12	9	10	22
Germany	2	1	2	3	4
Spain	3	5	11	9	17
Italy	4	19	16	23	37
Australia	5	11	12	15	14
UK	6	3	6	5	9
Japan	7	9	5	36	50
Korea, Rep.	8	8	4	61	30
Russia	9	22	29	22	25
Canada	10	32	28	51	54
Poland	11	6	3	1	3
China	12	2	1	2	1
Romania	19	4	7	7	5
Country	WInDed	WOutDeg	Eigenvector	LRIC (Max)	SRIC
India	29	7	8	4	6
Philippines	34	13	13	11	8
Morocco	35	14	17	8	7
Mexico	46	10	10	6	2
Algeria	75	50	61	41	10

2.1.6 Сетевой анализ международной торговли

Сетевой анализ международной торговли товарами проводился во многих исследованиях. Впервые подобная работа была опубликована в середине XX века и демонстрировала «… существенное единство мировой торговли» [151], однако большее распространение сетевой анализ получил в начале XXI века. Так, одно из исследований применяло различные меры сетевого анализа, такие как центральность по степени, центральность по близости, центральность по посредничеству для графического и аналитического изучения взаимосвязанности участников мировой торговли товарами и неоднородности мировой торговли товарами [152]. Впоследствии для исследования мировой торговли товарами было предложено визуализировать международную торговую сеть либо по географическому пространству, отражающему только прямые торговые потоки, либо по социограмме стран, отражающей прямые и косвен-

ные торговые потоки [153]. Также исследователи определяли наиболее влиятельных участников мировой торговли: в 2001 году наиболее значимыми в сети по международной торговле были Соединенные Штаты, Германия и Великобритания, а в 2010 году Соединенные Штаты, Китай и Германия [154].

Источником информации об экспортно-импортных операциях стала база статистических данных Организации Объединенных Наций по торговле товарами, являющаяся надежными источником информации о международной торговле товарами, поскольку собирается на основе данных, официально предоставленных Организации Объединенных Наций статистическими органами стран, зависимых территорий и территорий специального географического интереса. Собранные данные обладали следующими значимыми характеристиками: двусторонней асимметричностью, представляющей собой указание различной величины одного и того же торгового потока различными статистическими органами, и односторонней неполнотой, представляющей собой указание величины торгового потока только одним статистическим органом. В рамках данной работы с тем, чтобы рассчитать различные индексы центральности, использовались три базы данных, основанные на описанной выше выгрузке: экспортная база данных, импортная база данных и обработанная база данных.

Расчеты классических индексов центральности и новых индексов центральности продемонстрировали, что в период с 2001 по 2016 год существовало несколько лидеров мировой торговли: так, взвешенная центральность по степени продемонстрировала, что в 2001 году наиболее влиятельными странами были Соединенные Штаты, Германия, Япония, а в 2016 году Китай, Соединенные Штаты, Германия; PageRank центральность продемонстрировала, что в 2001 году наиболее влиятельными странами были Соединенные Штаты, Германия, Великобритания, а в 2016 – Соединенные Штаты, Китай, Германия; SRIC продемонстрировала, что 2001 году наиболее влиятельными странами были Соединенные Штаты, Германия, Япония, в 2016 году – Соединенные

Штаты, Китай, Германия; отличные результаты от предыдущих индексов центральности были характерны для LRIC: в 2001 году наиболее влиятельными странами были Испания, Южная Корея, Мексика, в 2016 – Испания, Россия, Канада. Таким образом, большинство индексов центральности в качестве наиболее влиятельных стран выделяют Соединенные Штаты, Германию и Китай.

2.1.7 Некоторые выводы

В рамках работы был предложен подход к оценке влияния в мультиплексных сетях. Наша модель ориентирована для случаев, когда нет четкой информации о зависимости между различными слоями. Важной частью нашей работы является преобразование сети, которое осуществляется с учетом неоднородности исходной сети. Ряд процедур коллективного выбора был адаптирован к оценке влияния в сетевых структурах. Наконец, предложенная модель была применена к сетям международной торговли продовольствием и торговли вооружением, а полученные результаты позволили выявить наиболее важных ее участников.

Предложенная модель имеет программную реализацию на языке Python и может быть скачана по адресу: https://github.com/SergSHV/slric.

Также был предложен ряд мер по оценке важности рёбер. Нашей целью было выявить те рёбра, через которые проходит более сильное влияние по сравнению с влиянием, которое проходит напрямую между вершинами графа. Ключевые рёбра, выявленные данным способом, помогают повысить общее влияние в сетевой структуре. При удалении важных рёбер общее влияние в графе ослабевает. Результаты данной работы были опубликованы в статье [155].

Также важно отметить, что существует много подходов к выделению наиболее влиятельных вершин в сети, обладающих различными свойствами и соответствующих различным наборам аксиом. Дальнейшим развитием обзор-

ной работы по анализу аксиоматики в графах является исследование соответствия мер большому пулу одинаковых аксиом с соответствующими ограничениями, с целью систематизации и классификации мер центральности. Унифицированный подход к выбору методов может упростить задачу анализа сетей.

В работе по анализу миграционных потоков проведен анализ международной миграции с 1970 по 2014гг. на основе данных по потокам миграции между странами. С помощью сетевого анализа определены рейтинги влиятельных стран в процессе международной миграции с помощью индексов центральностей, как классических, так и учитывающих непрямые взаимодействия. Также собраны данные по потокам международной миграции и другим страновым показателям за 2015-2017гг. Данные результаты будут использованы для дальнейшего исследования сетевого анализа международной миграции.

Исследовательская работа по анализу международной торговли была направлена на одновременный анализ прямого и непрямого взаимодействия между странами посредством применения классических индексов центральности и новых индексов центральности. При этом, данные центральности были рассчитаны на нескольких базах данных: обработанной базе данных, импортной базе данных, экспортной базе данных (во многих случаях результаты, рассчитанные на основе обработанной базы данных, были аналогичны результатам, рассчитанным на основе базы данных импорта, чем на базе данных экспорта, что обусловлено большим количеством торговых потоков в базе данных импорта, чем в базе данных экспорта). Несмотря на то, что применялись различные индексы центральности и, следовательно, различные аспекты мировой торговли принимались как важные, три страны были лидерами мировой торговли: Соединенные Штаты, Германия и Китай. Данные результаты могут быть применены исследовательскими центрами, занимающимися обработкой международных торговых данных для исключения в них двусторонней асимметрии и односторонней неполноты (например, Центром перспективных ис-

следований и международной информации), а также международными организациями, исследующими мировую торговлю (например, Организацией Объединенных Наций, Международным торговым центром, Европейским статистическим бюро по международной торговле товарами и услугами).

2.2 Исследования по анализу текстов и кластерному анализу

2.2.1 Введение и постановка задачи оптимального обобщения нечетких множеств

Проблемы автоматизации структурирования и интерпретации результатов поиска в коллекциях текстовых документов приобретает все большее значение как из-за практических потребностей, так и по причине теоретической необходимости. Для облегчения интерпретации коллекций разнородных документов, получаемых в ответ на те или иные запросы, фундаментальный характер приобретает необходимость сокращенного описания результатов поиска. Современные методы сокращения текстов основаны на “механистическом” вычленении тех или иных фрагментов текстов для помещения их в итоговое сокращенное представление. По нашему мнению, решение проблемы невозможно без попыток автоматизации когнитивной способности к обобщению, то есть замене совокупности “частных” понятий небольшим количеством обобщающих концепций. Это соответствует хорошо знакомому смыслу термина “обобщать” (см., например, Современный толковый словарь русского языка Ефремовой: “обобщать” – делать общим, соединять вместе; объединять).

Существующие подходы к вычислительному анализу структуры текстовых коллекций обычно не предполагают обобщения в качестве конкретно поставленной цели. Наиболее популярными инструментами для структурирования текстовых коллекций являются кластерный анализ и тематическое моделирование. Оба способа включают признаки того же уровня детализации, что и отдельные слова или короткие фразы в текстах, таким

образом, не давая решения задачи обобщения в качестве четко сформулированной цели.

Тем не менее, иерархическая природа значений понятий и терминов отражается в потоке публикаций по анализу текста. Мы можем различать по крайней мере три направления, в контексте которых рассматривается вопрос обобщения. Прежде всего, следует упомянуть деятельность, связанную с развитием таксономий, в особенности, гипонимические / гипернимические отношения (см., например, [156, 157], и ссылки в них). Также следует упомянуть недавнюю статью [158], посвященную дополнению таксономии новыми темами исследований.

Другое направление является разделом классической деятельности по обобщению текста - автоматическое “аннотирование”, “реферирование” или “резюмирование”, позволяющее, в частности, получать краткие “выжимки” из более объемных текстов. Обычно, на практике “резюме” текста создаются с использованием механистических подходов к извлечению предложений. Тем не менее, отметим также подход для построения резюме как абстракций текстов путем объединения некоторых шаблонов, таких как триплеты субъект-глагол-объект (SVO) (см., например, [159]).

Еще одна сфера деятельности – это то, что можно назвать “операционным обобщением”. В этом направлении авторы используют обобщенные описания с использованием таксономических отношений между обобщенными состояниями и их составными частями для достижения определенных целей, таких как улучшение характеристик поиска текста (см., например, [160, 161]).

Мы будем использовать существующие таксономии для реализации идеи обобщения, не склоняясь ни к какому из описанных выше направлений. В соответствии со значением из Словаря современного русского языка, приведенным выше, будем понимать термин “обобщение” в смысле вывода общей концепции из отдельных частностей. Мы предполагаем, что нам дается самая простая структура для такого вывода – таксономия предметной области.

В этом понимании таксономия представляет собой корневое дерево над ключевыми понятиями предметной области, помечающими узлы дерева. Дерево отражает основные гипонимические/гипернимические отношения в области так, что узел A является родителем узла B, если отношение “B входит в A” истинно. Ситуация, которая требует особенного внимания – это случай, когда мы должны обобщить нечеткий набор листов таксономии, представляя таким образом суть какого-то эмпирически наблюдаемого явления.

В частности, рассмотрим моделирование направлений новых исследований в области Науки о данных (Data Science). Data Science – это новая область компьютерных наук. Самая популярная таксономия в области компьютерных наук разработана вручную Всемирной ассоциацией вычислительной техники (Association for Computing Machinery – ACM), наиболее представительным органом в этой области; последний выпуск таксономии был опубликован в 2012 году под наименованием “Система классификации компьютерных наук ACM” (ACM Computing Classification System – ACM-CCS) [ACM 2012]. Мы рассматриваем ее часть, связанную с Наукой о данных, расширенную с помощью добавления нескольких листьев (описанную в [162]). Добавление листьев сделано для того, чтобы лучше отразить тематику анализируемых научных статей рассматриваемой в дальнейшем коллекции [163].

В качестве коллекции документов мы рассмотрим коллекцию научных статей, опубликованных за последний период в ряде журналов, представляющих область Data Science (Наука о данных).

На примере этой коллекции и таксономии ACM-CCS мы собираемся найти кластеры тем максимально “тесным” образом, каждый из которых представляет тенденцию развития, отраженную в коллекции. В этом случае естественно предположить, что кластеры будут нечеткими, в соответствии с нечеткой природой семантики научных статей. Основной проблемой является интерпретация такого нечеткого кластера. Как уже указывалось, обычно интерпретация кластеров осуществляется с использованием понятий того же

уровня гранулярности (признаки, ключевые слова), которые используются при построении кластеров. Мы предлагаем использовать понятия укрупненной гранулярности, для чего воспользоваться таксономией предметной области. Конкретно, речь идет об обобщении понятий таксономии, соответствующих полученному кластеру.

2.2.2 Экономный подъем нечеткого тематического кластера в таксономии: модель и метод

Говоря математическим языком, таксономия – это корневое дерево, чьи узлы аннотированы темами таксономии. Мы будем рассматривать следующую задачу: дано нечеткое множество S из листьев таксономии, необходимо найти узел $t(S)$ более высокого ранга в таксономии, который покрывает множество S наиболее “тесным” образом. Такая задача “подъема” – математическая постановка человеческой способности делать обобщение, то есть, процесса формирования “концептуальной” формы представленного явления, в данном случае, нечеткого множества листьев дерева.

Задача не так проста, как могло бы показаться. Для краткости, рассмотрим четкое множество S, представленное пятью черными прямоугольниками на фрагменте дерева на рисунке 2.5. Рисунок 2.6 иллюстрирует ситуацию, при которой множество поднято в корень, что показано закрашиванием соответствующего ему прямоугольника, как и его потомков.

Если мы допустим, что множество S может быть обобщено корнем, это приведет к тому, что корнем будет покрыто и четыре белых прямоугольника, и, таким образом, попадет в то же обобщение, что и S, при том, что они не принадлежат множеству S. Такую ситуацию будем далее называть появлением пробелов. Очевидно, что вариант подъема, при котором возникают пробелы, должен штрафоваться моделью подъема. В целом, число концептуальных элементов, введенных для обобщения S, здесь равно 1 головной вершине (то есть корню), который мы теперь присвоили переменной S, и 4 пробела,

образовавшиеся из-за топологии дерева. При этом за подобный подъем, как уже было сказано, должно быть назначено определенное значение штрафа. Другой возможный вариант подъема показан на рисунке 2.7: здесь заданное множество поднимается только до корня левой ветви дерева. Мы можем видеть, что число пробелов резко сократилось – до 1. Однако возник еще один тип неточности подъема: черный прямоугольник справа, принадлежащий *S*, не покрыт корнем левой ветви, в которую отображается множество *S*. Этот тип ошибки будем называть выбросом. При этом варианте подъема появляются три новых объекта: одна головная вершина, один выброс и один пробел. Это меньше, чем количество объектов, появившихся при поднятии множества к корню (одна головная вершина и четыре пробела, то есть совокупно – пять), что делает его более предпочтительным. Конечно, этот вывод верен только в том случае, если относительный штраф за один выброс меньше суммарного штрафа за три пробела.

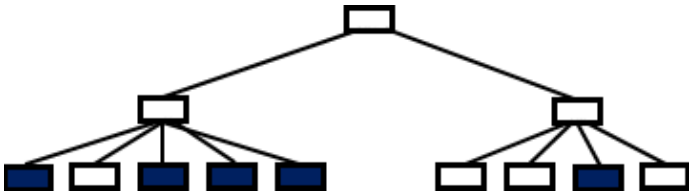


Рисунок 2.5 – Фрагмент таксономии, черные прямоугольники соответствуют элементам, принадлежащим множеству

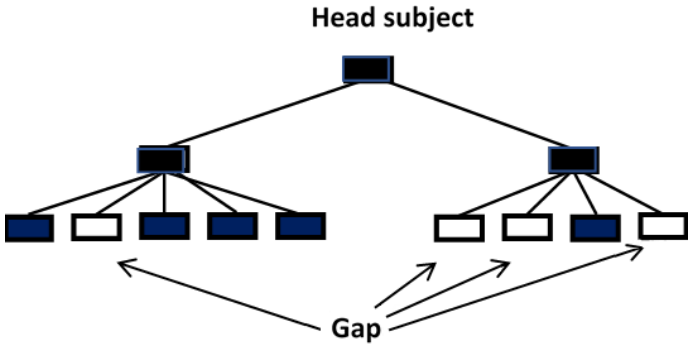


Рисунок 2.6 – Обобщение нечеткого множества запроса, показанного на рисунке 2.5 с помощью его отображения в корень дерева. Цена этого обобщения – четыре пробела (gap)

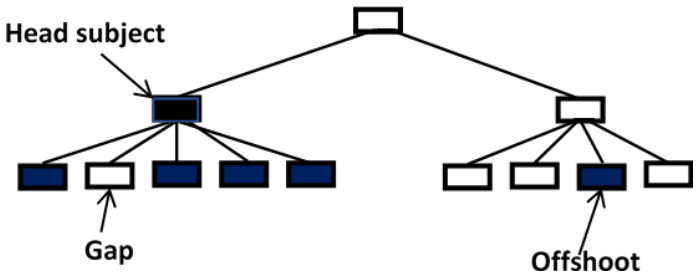


Рисунок 2.7 – Обобщение нечеткого множества запроса, показанного на рисунке 2.5 с помощью его отображения в корень левого поддерева. Цена такого обобщения – один пробел (gap) и один выброс (offshoot)

Нас интересует, можно ли обобщить нечеткое множество *S* в узел *t* принадлежащий более высокому рангу таксономии, так, чтобы можно было считать, что *S* попадает в покрытие, охватываемое узлом *t*. Если это будет сделано напрямую, например, путем выбора последнего общего предка для узлов в *S* получение желаемого охвата будет маловероятным. Чтобы сделать процесс обобщения более адекватным, следует допустить наличие в итоговом обобщении не только пробелов, но и выбросов, являющихся листьями, принадлежащими обобщению, но не исходному множеству.

Цель нахождения интерпретируемого обобщения для *S* в рамках таксономии может быть формализована как нахождение одного или нескольких “головных тем” для покрытия *S* с минимальным количеством новых элементов всех типов, получившихся в процессе обобщения: главных вершин, пробелов и выбросов. Такая постановка цели реализует принцип максимальной экономии (Maximum Parsimony, MP) при описании данного явления, популярный в некоторых областях, таких как биоинформатика.

Рассмотрим корневое дерево *T*, представляющее иерархическую таксономию, такую, что его узлы снабжены ключевыми фразами, обозначающими различные понятия. Обозначим множество его листьев через *I*. Соотношение между узлами в иерархии обычно выражается с использованием генеалогических терминов: каждый узел $t \in T$ называется родителем узлов, непосредственно ответвляющихся от *t*, его потомков. Будем использовать $\chi(t)$ для обозначения множества потомков этого узла.

Предполагается, что каждый внутренний узел $t \in T - I$, то есть узел, не являющийся листом таксономии, соответствует понятию, обобщающему темы, соответствующие листьям *I(t)*, выходящим из *t*, то есть, листья поддерева *T(t)* имеют общий корень в узле *t*, называющийся кластером листьев.

Нечеткое множество на I – это отображение u из множества I в множество неотрицательных действительных чисел, которое присваивает значение функции принадлежности (или поддержку) $u(i) \geq 0$ каждому узлу $i \in I$. Мы называем множество $S_u \subset I$, где $S_u = \{i \in I: u(i) > 0\}$, поддержкой u. Никаких других предположений о виде функции u не делается. Отметим только то, что для удобства целесообразно положить такое ограничение, чтобы ее значение не превышало единицу. Классические, или четкие множества соответствуют бинарным функциям принадлежности – таким, что *u(i)*=1, если $i \in S_u$ и *u(i)*=0 иначе.

Рассматривая нечеткое множество запроса *u*, определенное на листьях *I* дерева *T*, можно считать *u* (возможно, сделанной с погрешностями) проекцией понятия более высокого ранга, “головной темы”, на соответствующий кластер листьев. В этом предположении должен существовать головной узел *h* среди внутренних узлов дерева *T* такой, что его кластер листьев *I(h)* в определенной степени совпадает (с небольшими ошибками) с S_u. Этот головной узел – обобщение *u*, которое и необходимо найти.

Возможны два типа ошибок, связанных с поиском головного узла, если он точно не покрывает поддержку, которые представляют собой ложные срабатывания и ложные отрицания, уже упомянутые ранее как пробелы (gaps) и выбросы (offshoots) – соответственно, иллюстрируются на рисунках 2.6 и 2.7. Общее количество головных узлов, пробелов и выбросов должно быть как можно меньше. Для этого мы вводим штраф за каждый из этих элементов. Предполагая, для простоты, что все черные прямоугольники на рисунке 2.5 имеют значения функции принадлежности, равные единице, можно легко увидеть, например, что общее значение штрафа, например, для подъема на

рисунке 2.7 равно $1 + \gamma + \lambda$, где γ, λ и 1 – штрафы за выброс, пробел и головную тему соответственно.

Рассмотрим узел-кандидат h в T и его значение принадлежности в нечетком множестве u. h-пробел – это узел g из $T(h)$, отличный от h, в котором произошла потеря значения, то есть g является максимальным u-иррелевантным узлом в том смысле, что его родитель не является u-иррелевантным. И наоборот, установление узла h в качестве головного можно рассматривать как приобретение значения u в узле. Множество всех h-выбросов обозначим как $G(h)$.

Чтобы правильно определить идею использования пробелов, введем понятие u-иррелевантных узлов. Узел t называется u-иррелевантным, если его кластер листьев $I(t)$ не пересекается с поддержкой S_u. Очевидно, что если узел принадлежит u, все его потомки также принадлежат u. h-выброс – это лист $i \in S_u$, который не покрывается h, то есть $i \notin I(h)$. Множество всех h-выбросов – это $S_u - I(h)$.

Если задано нечеткое тематическое множество u над I, множество узлов H будет называться u-покрытием, если: (a) H покрывает S_u, то есть, $S_u \subseteq \bigcup_{h \in H} I(h)$, и (b) узлы в H не связаны, то есть $I(h) \cap I(h') = \emptyset$ для любых $h, h' \in H$, таких, что $h \neq h'$. Внутренние узлы H будут называться головными темами, а листовые узлы – выбросами, поэтому множество выбросов в H равно $H \cap I$. Множество пробелов в H является объединением $G(h)$ по всем головным темам $h \in H - I$.

Определим штрафную функцию $p(H)$ для u-покрытия H следующим образом:

$$p(H) = \sum_{h \in H-I} u(h) + \sum_{h \in H-I} \sum_{g \in G(h)} \lambda v(g) + \sum_{h \in H \cap I} \gamma u(h).$$

Задача, которую мы решаем, состоит в том, чтобы найти такое u-покрытие H, которое минимизирует штрафную функцию $p(H)$. Такое u-покрытие и будет наиболее оптимальным обобщением нечеткого множества u.

Перед применением алгоритма для минимизации суммарного штрафа необходимо выполнить предварительное преобразование дерева. Для этого необходимо сделать следующие шаги. Из дерева нужно удалить все немаксимальные u-иррелевантные узлы, то есть потомки пробелов, а также найти множества *G(t)* и значения суммарной важности пробелов $V(t) = \sum_{g \in G(t)} v(g)$ для всех внутренних узлов *t*. Отметим, что все элементы S_u находятся в множестве листьев обрезанного дерева, а остальные листья обрезанного дерева – это именно пробелы.

Теперь мы можем применить алгоритм подъема ПарГеНМ. Для каждого узла *t* алгоритм ПарГеНМ вычисляет два набора – *H(t)* и *L(t)*, содержащие такие узлы в *T(t)*, в которых, соответственно, происходят приобретения и потери головных узлов (включая выбросы). Соответствующий штраф обозначается *p(t)* и вычисляется так, как описано ниже.

Предположение алгоритма состоит в том, что после потери узла не может быть последующего ее приобретения. Следовательно, *H(t)* и *L(t)* вычисляются при условии, что головной узел не был ни приобретен (ни потерян) ни у одного из предков *t*. Алгоритм ПарГеНМ рекурсивно вычисляет *H(t)*, *L(t)* и *p(t)* из соответствующих значений для дочерних узлов в $\chi(t)$.

Для каждого листового узла, который не принадлежит S_u, мы определяем как *L(.)* так и *H(.)* пустыми, а штраф равным нулю. Для каждого листового узла, который находится в S_u, *L(.)* устанавливается пустым, тогда как *H(.)*, содержит только листовой узел, и штраф определяется как произведение зафиксированного параметра γ на значение его принадлежности. Чтобы вычислить *L(t)* и *H(t)* для любого внутреннего узла *t*, необходимо определить, какая из двух альтернатив имела место:

(a) головной узел был приобретен в *t*;

(b) головной узел не был приобретен в *t*.

В случае (a) множества *H(.)* и *L(.)* у его потомков далее оказываются не нужными. В этом случае, *H(t)*, *L(t)* и *p(t)* определяются как:

$$\begin{cases} H(t) = \{t\}, \\ L(t) = G(t), \\ p(t) = u(t) + \lambda V(t). \end{cases}$$

В случае (b) множества *H(t)* и *L(t)* являются просто объединениями соответствующих множеств из их потомков, а *p(t)* – сумма всех штрафов из потомков:

$$\begin{cases} H(t) = \bigcup_{w \in \chi(t)} H(w), \\ L(t) = \bigcup_{w \in \chi(t)} L(w), \\ p(t) = \sum_{w \in \chi(t)} p(w). \end{cases}$$

Чтобы получить оптимальный подъем, из этих случаев выбирается тот, где достигается меньшее значение *p(t)*. В ситуации, когда оба случая дают одинаковые значения *p(t)*, мы можем выбрать любой из них произвольно – в формулировке алгоритма ниже, мы выбрали (a). Выходные данные алгоритма включают значения рассчитываемых множеств в корне, а именно: H – множество головных узлов и выбросов, L – множество пробелов и p – назначенное значение штрафа. Мы доказали, что алгоритм ПарГеНМ действительно приводит к глобально оптимальному обобщению [163].

2.2.3 Структурирование коллекции научных публикаций с помощью таксономии предметной области

Продемонстрируем применение алгоритма ПарГеНМ для задачи анализа коллекций на примере работы с коллекцией научных публикаций.

Столкнувшись с проблемой структурирования и интерпретации множества научных публикаций в заданной предметной области, можно предложить два способа анализа: первый – эмпирический, второй – теоретический. Первый способ предлагает определить категории,

присутствующие в коллекции, из самих текстов, второй – на основе знаний о предметной области. Первый подход, например, может быть хорошо проиллюстрирован тематическим моделированием на основе LDA (в частности, [164]); второй подход – применением таксономии, разрабатываемой экспертами, такой как ACM-CCS (подход описан, например, в [165]).

Для анализа коллекции мы последовательно выполним следующие шаги:

– подготовка коллекции научных статей;

– подготовка таксономии рассматриваемой предметной области;

– расчет матрицы значений релевантности между тематическими листьями таксономии и публикациями из коллекции статей;

– поиск нечетких кластеров по матрице значений релевантности;

– подъем полученных кластеров в таксономии для их концептуализации с помощью обобщения;

– построение выводов из обобщений.

Каждый из этих пунктов далее рассматривается в отдельном подразделе.

2.2.3.1 Подготовка коллекции научных публикаций

Благодаря доступу к библиотеке журналов издательства Springer нам удалось загрузить коллекцию из 17685 научных статей вместе с их аннотациями, опубликованных в 17 журналах, связанных с Наукой о данных, за 20 лет: с 1998 по 2017 год. В качестве текстовой коллекции были рассмотрены аннотации этих статей.

2.2.3.2 Таксономия DST

Таксономия – это форма инженерии знаний, которая в последние годы становится все более популярной. Наиболее известны таксономии, созданные

в рамках проекта Genome Ontology (GO) [166] в области биоинформатики, проекта SNOMED CT для здравоохранения и медицины [167] и тому подобное.

Математически таксономия – это корневое дерево, иерархия, все узлы которой помечены основными понятиями предметной области. Иерархия соответствует отношению включения: тот факт, что узел A является родителем B, означает, что B является частью или частным случаем A. Важной характеристикой корневого дерева является то, что у каждого узла есть только один родительский элемент.

Существует два основных подхода к разработке таксономии домена: автоматический и ручной. Последний, безусловно, в настоящее время является гораздо более зрелым и развитым. Тем не менее, даже этот подход страдает от недостатков, которые сводятся к тому, что решения по разработке таксономии, касающиеся используемых структур классификации, процедур и описательных основ, как правило, недостаточно хорошо описаны и мотивированы. Автоматический подход может использовать множество цифровых ресурсов и методов семантического анализа. На данный момент чаще предпочитаются таксономии, разрабатываемые вручную представительным органом специалистов. Действительно, такая таксономия не зависит от чисто эмпирических данных, используемых автоматическими методами.

Более того, ручная таксономия обычно уравновешивает теоретические наработки и практический опыт, накопленный в сообществе, представленном органом, который ее разрабатывает. Такова, например, Система классификации компьютерных наук Всемирной ассоциации вычислительной техники ACM (ACM-CCS 2012). Выбранное нами ее подмножество Data Science – Наука о данных, включает такие области, как машинное обучение, интеллектуальный анализ данных, анализ данных и так далее. Мы берем ту часть таксономии ACM-CCS 2012, которая связана с Data Science, и добавляем несколько листов, связанных с более поздними разработками Data Science.

Рассматриваемая таксономия целиком опубликована в [162]. Темы высших рангов таксономии представлены в Таблице 2.6.

Таблица 2.6 – Подмножество таксономии Системы классификации компьютерных наук Всемирной ассоциации вычислительной техники ACM (ACM-CCS) 2012, касающееся Науки о данных; несколько тем самых высоких рангов.

Subject index	Subject name
1.	Theory of computation
1.1.	Theory and algorithms for application domains
2.	Mathematics of computing
2.1.	Probability and statistics
3.	Information systems
3.1.	Data management systems
3.2.	Information systems applications
3.3.	World Wide Web
3.4.	Information retrieval
4.	Human-centered computing
4.1.	Visualization
5.	Computing methodologies
5.1.	Artificial intelligence
5.2.	Machine learning

2.2.3.3 Вычисление степени релевантности между текстами и темами таксономии

Для вычисления степеней релевантности тем таксономии и текстовых документов коллекции будем использовать метод АСД, описанный, например, в [168]. Как уже было сказано, существует большое количество его валидаций на различных коллекциях, показавших его эффективность. Данный метод не требует предобработки текстов и учитывает порядок слов, в отличие от большинства популярных подходов.

Чтобы уменьшить влияние часто встречающихся общих терминов, функция оценки модифицируется введением пятикратного уменьшения веса специальных стоп-слов. Список стоп-слов включает: “learning, analysis, data, method” и несколько постфиксов: “s/es, ing, tion”.

2.2.3.4 Определение и вычисление нечетких кластеров тем таксономии

Группы тем, которые мы планируем получить, должны отражать одновременное возникновение тем: чем больше число текстов, к которым относятся обе темы t и t', тем больше взаимосвязь между t и t', тем больше вероятность того, что темы t и t' попадут в один кластер.

Было протестировано несколько популярных алгоритмов кластеризации: K-Means, DBSCAN и классический спектральный метод, которые не показали удовлетворительных результатов, ввиду того, что не дали интерпретируемых кластеров. Поэтому мы представляем здесь результаты, полученные с помощью алгоритма FADDIS, описанного в [165], и разработанного специально для поиска нечетких тематических кластеров. Этот алгоритм реализует следующие предположения, относящиеся к рассматриваемой задаче:

– [LN] нормализация Лапласа: преобразование матрицы похожести, выполняемое для того, чтобы сделать кластерную структуру четче;
– [AA] аддитивность: тематические кластеры за текстами являются аддитивными, поэтому значения сходства представляют собой суммы вкладов по различным скрытым темам;
– [AN] неполнота: кластеры не обязательно охватывают все ключевые фразы, имеющиеся в наличии, поскольку рассматриваемая текстовая коллекция может не иметь отношения к некоторым из них.

2.2.3.4.1 Оценивание степени близости тем

Основываясь на матрице релевантности ключевых фраз и документов коллекции R, можно получить матрицу близости между ключевыми фразами – A, которая, говоря иными словами, представляет собой матрицу “совместной релевантности” ключевых фраз в соответствии со структурой текстовой коллекции. Элементы этой матрицы, то есть значения сходства $a_{tt'}$ между темами t и t' можно вычислить как скалярное произведение векторов оценок

$r_t = (r_{tv})$ и $r_{t'} = (r_{t'v})$, где v=1,2,...,V=17685. Скалярное произведение можно модифицировать, используя естественные весовые коэффициенты, присваиваемые документам в коллекции. Вес текстового документа v определяется как отношение количества всех относящихся к нему тем n_v и n_{max}, максимуму n_v для всех v = 1,2,...,V. Тема считается относящейся к v, если показатель ее релевантности документу превышает 0.2 (этот порог найден экспериментально в [168]).

2.2.3.4.2 Аддитивная нечеткая спектральная кластеризация

Обозначим множество листовых тем через T и предположим, что нечеткий кластер над T представлен вектором принадлежности $\vec{u} = u_t$, $t \in T$, таким что $0 \leq u_t \leq 1$ для всех $t \in T$, а также пусть интенсивность $\mu > 0$ – масштабирующий коэффициент, связывающий значения принадлежности со значениями близости. Для T, представляющего собой набор тем исследований, и $\vec{u} = u_t$, $t \in T$, вектора значений принадлежности, представляющего собой семантическую подструктуру совокупности рассматриваемых публикаций, произведение $(\mu u_t)(\mu u_{t'}) = \mu^2$ можно рассматривать как вклад направления исследований, представленного рассматриваемым кластером в общую оценку близости $a_{tt'}$ между темами t и t'.

Модель аддитивной нечеткой кластеризации, предложенная в [169], утверждает, что значения элементов матрицы близости тем A можно рассматривать как результат аддитивного вклада нечетких кластеров K, вплоть до ошибок, которые должны быть минимизированы:

$$a_{tt'} = \sum_{k=1}^{K} \mu_k^2 \, u_{kt} u_{kt'} + e_{tt'}.$$

где $\overrightarrow{u_k} = (u_{kt})$ – вектор принадлежности кластера k, и μ_k – его интенсивность.

Эти предположения требуют, чтобы кластеры были извлечены в соответствии с аддитивной моделью. Метод, разработанный в [169], FADDIS, находит кластеры последовательно, один за другим, что соответствует сформулированным выше предположениям, минимизируя:

$$E = \sum_{t,t' \in T} (w_{tt'} - \chi u_t u_{t'})^2.$$

В [169] приведено большое количество теоретических и экспериментальных результатов, призванных продемонстрировать тот факт, что FADDIS конкурентоспособен по сравнению с другими подходами нечеткой кластеризации и имеет ряд преимуществ.

2.2.3.4.3 Тематические кластеры FADDIS

После вычисления матрицы близости тем размера 317×317 (по количеству листовых тем таксономии), применения псевдообратного преобразования Лапласа, и кластеризации методом FADDIS, мы последовательно получили 6 кластеров, три из которых кажутся особенно однородными. Основываясь на их содержимом, 'Learning', 'Retrieval' и 'Clustering', обозначим их L, R и C соответственно. Эти кластеры представлены в Таблице 2.7.

Таблица 2.7 – Кластеры L, R, C: темы с наибольшими значениями функции принадлежности

Cluster L			Cluster R			Cluster C		
$u(t)$	Code	Topic	$u(t)$	Code	Topic	$u(t)$	Code	Topic
0.300	5.2.3.8.	rule learning	0.211	3.4.2.1.	query representation	0.327	3.2.1.4.7.	biclustering
0.282	5.2.2.1.	batch learning	0.207	5.1.3.2.1.	image representations	0.286	3.2.1.4.3.	fuzzy clust.
0.276	5.2.1.1.2.	learning to rank	0.194	5.1.3.2.2.	shape representations	0.248	3.2.1.4.2.	consensus clust.
0.217	1.1.1.11.	query learning	0.194	5.2.3.6.2.1.	tensor representation	0.220	3.2.1.4.6.	conceptual clust.
0.216	5.2.1.3.3.	apprenticeship learning	0.191	5.2.3.3.3.2.	fuzzy representation	0.192	5.2.4.3.1.	spectral clust.
0.213	1.1.1.10.	models of learning	0.187	3.1.1.5.3.	data provenance	0.187	3.2.1.4.1.	massive data clust.
0.203	5.2.1.3.5.	Adversarial learning	0.173	2.1.1.5.	equational models	0.159	3.2.1.7.3.	graph based conceptual clust.

2.2.3.5 Результаты подъема кластеров L, R, и C в таксономии DST

Полученные кластеры были подняты в таксономии DST с использованием алгоритма ПарГеНМ, со штрафом за пробел λ=0.1 и выброс γ=0.9.

Результаты подъема кластера L показаны на рисунке 2.8. Кластер получил три головных темы: machine learning, machine learning theory, and learning to rank. Они представляют структуру общего понимания “Learning” в соответствии с нашей текстовой коллекцией. Список полученных пробелов является менее представительным, отражая, вероятно, относительно скромный охват этих областей публикациями в коллекции (пробелы показаны в Таблице 2.8).

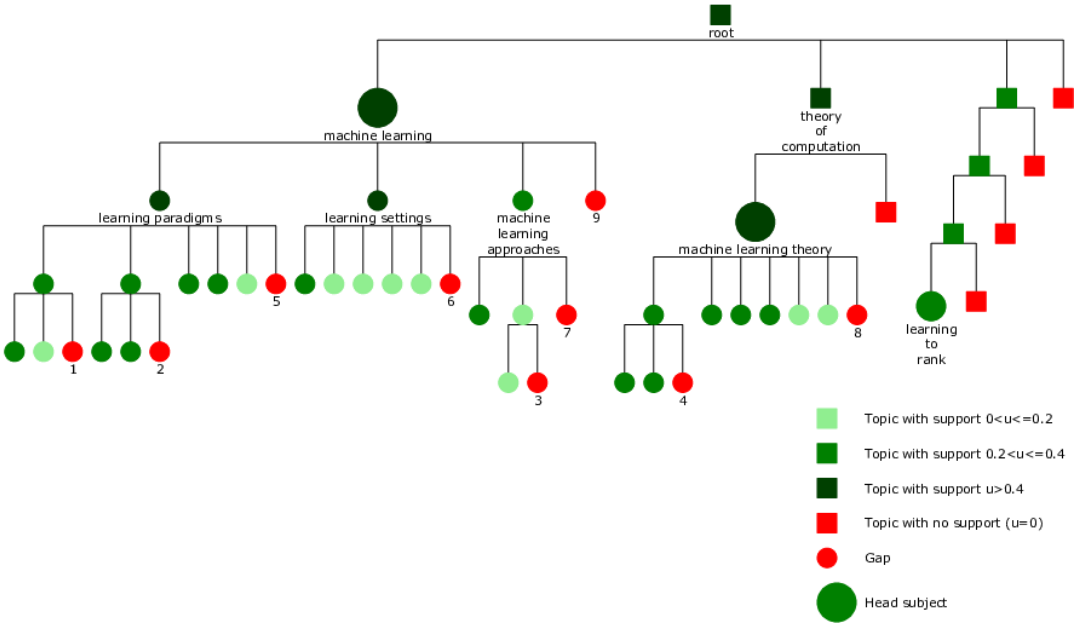


Рисунок 2.8 – Результаты подъема кластера L: Learning. Пробелы пронумерованы, см. Таблицу 2.8

Таблица 2.8 – Пробелы, полученные в процессе подъема кластера L

Номер	Темы
1	ranking, supervised learning by classification, structured outputs
2	sequential decision making in practice, inverse reinforcement learning in practice
3	statistical relational learning
4	sequential decision making, inverse reinforcement learning
5	unsupervised learning
6	learning from demonstrations, kernel approach
7	classification and regression trees, kernel methods, neural networks, learning in probabilistic graphical models, learning linear models, factorization methods, markov decision processes, stochastic games, learning latent representations, multiresolution, support vector machines
8	sample complexity and generalization bounds, boolean function learning, kernel methods, boosting, bayesian analysis, inductive inference, structured prediction, markov decision processes, regret bounds
9	machine learning algorithms

Аналогичные комментарии можно сделать и в отношении результатов подъема кластера R: Retrieval (рисунок 2.9). Получены следующие головные темы: Information Systems and Computer Vision, они показывают структуру публикаций по тематике “Retrieval” в рассматриваемой коллекции.

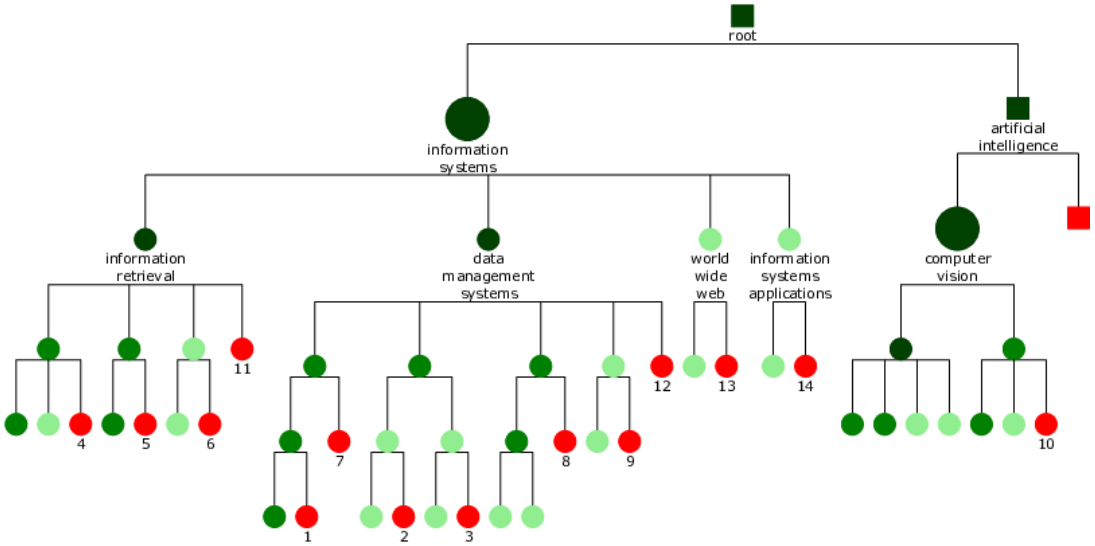


Рисунок 2.9 – Результаты подъема кластера R. Пробелы пронумерованы. Обозначения к рисунку приведены в легенде рисунка 2.8

Мы не приводим иллюстрции подъема кластера C, так как соответствующий фрагмент таксономии оказался слишком большим, а результаты – слишком фрагментарны. Здесь было получено 16 (!) головных тем: clustering, graph based conceptual clustering, trajectory clustering, clustering and classification, unsupervised learning and clustering, spectral methods, document filtering, language models, music retrieval, collaborative search, database views, stream management, database recovery, mapreduce languages, logic and databases, language resources. Как можно видеть, ключевые темы кластеризации дополняются методами и приложениями – это показывает, что постоянно возрастающая роль кластеризации в Науке о данных, возможно, должна быть лучше отражена в таксономии.

2.2.4 Некоторые выводы

Можно видеть, что тематические кластеры, найденные в рассматриваемой коллекции научных публикаций, составляют области наиболее актуальных разработок. Три рассматриваемых кластера тесно связаны со следующими процессами:

– теоретические и методологические исследования в машинном обучении, а также слияние машинного обучения с задачами ранжирования;

– представление различных типов данных для поиска информации и их объединение с визуальными данными, а также их семантикой;

– различные типы кластеризации в разных отраслях таксономии, связанные с разнообразными приложениями.

В частности, из головных тем кластера “Learning” хорошо заметно (рисунок 2.8 и комментарии к нему), что основная деятельность здесь все еще сосредоточена на теории и методах, а не на приложениях. Хорошей новостью является то, что области машинного обучения, ранее сфокусированные в основном на задачах обучения на подмножествах и разбиениях, в настоящее время расширяются в направлении обучения ранжированию и рейтингованию. Конечно, остается много областей, которые еще предстоит исследовать более подробно: их можно увидеть в списке пробелов в Таблице 2.8.

Переходя к анализу результатов поднятия кластера R “Retrieval” (рисунок 2.9), мы можем четко видеть тенденции современного этапа развития этого раздела Науки о данных. Вместо того, чтобы относить термин “информация” только к текстам, как это было на предыдущих этапах развития информационных наук, в наши дни анализ визуальных представлений становится частью концепции информации. Однако есть одна загвоздка. В отличие от многоуровневой гранулярности представлений текстов, разработанной человечеством в течение тысячелетий в процессе общения через естественные языки, не существует сопоставимой иерархии для

изображений. Можно предположить, что элементы кластера R, относящиеся к сегментации изображений и видео, а также элементы, относящиеся к системам управления данными – это те элементы, которые будут положены в основу будущей многоуровневой системы анализа изображений и видео. Это – потенциальное направление для будущих разработок, оно совершенно четко видно из рисунка 2.9.

Что же касается кластера C “Clustering” с его 16 головными темами, можно сделать вывод, что, возможно, уже наступил тот момент времени (или же такой момент должен наступить уже очень скоро), когда тема кластеризации должна быть поднята на более высокий уровень таксономии, чтобы охватить все эти головные узлы. В начале века Data Science, несколько десятилетий назад, кластеризация, как правило, считалась в значительной степени вспомогательной частью машинного обучения, обучением без учителя. Возможно, вскоре мы увидим новую таксономию Data Science, в которой кластеризация – это не просто вспомогательный инструмент, а, скорее, модель эмпирической классификации, значительная часть инженерии знаний. Обсуждая роль классификации как явления инженерии знаний, мы сталкиваемся с тремя общепринятыми аспектами классификации:

– структурирование явлений;

– связывание различных аспектов явлений друг с другом;

– формирование и хранение знаний о явлениях.

Каждый из них может представлять отдельное направление исследований в области инженерии знаний.

Результаты работы были опубликованы в статьях [170, 171, 172, 173, 174, 175].

2.3 Разработка методов анализа паттернов

2.3.1 Введение

Необходимость анализа многофакторных объектов в таких сложно формализуемых областях как социология, политология, экономика и др. способствовало активному развитию инструментария и отдельных направлений Visual Data Mining. Одним из таких, уже зарекомендовавшим себя направлений, является визуализация данных с использованием параллельных координат. Укажем его основную идею. Пусть имеется некоторое множество исследуемых объектов. Каждый объект характеризуется набором m параметров, что будем записывать в векторной форме как $\left(p_{i1}, p_{i2}, \ldots, p_{ij}, \ldots, p_{im}\right)$, где p_{ij} – значение j-ого параметра i-ого объекта. С целью анализа данных, используется графическое представление объекта в системе параллельных координат [176, 177]. Данное представление использует m параллельных, обычно вертикальных и равномерно распределённых координатных осей, каждая из которых отражает один из выбранных параметров. Объект изображается в виде кусочно-линейных функций с вершинами на координатных осях. Такое графическое представление данных позволяет проводить их визуальный анализ, и в частности, объединять объекты в отдельные группы (кластеры), исходя из ко-монотонности их графиков [177]. Этот метод успешно зарекомендовал себя при решении прикладных задач макроэкономики, банковской сферы, политологии [178, 179, 180], исследовании распределении рабочего времени российских экономистов [181] Однако, следует отметить, что вид графиков существенным образом зависит от порядка расположения координатных осей, что в ряде случаев, приводит к неоднозначности результатов кластеризации [182].

В настоящей работе нас интересовало поиск кластеров, обладающих свойствами взаимной симметрии кусочно-линейных функций, которая не должна нарушаться с изменением порядка расположения координатных осей.

Поэтому, далее, мы будем опираться на метод анализа паттернов, обеспечивающий это свойство. Нас будет интересовать также смысловое содержание, отражаемое в каждом отдельном случае свойством взаимной симметрии.

2.3.2 Порядково-инвариантная паттерн-кластеризация

Основная идея метода порядково-инвариантной паттерн-кластеризации состоит в поиске и объединении объектов, которые не меняют свою принадлежность группе (т.е. сохраняют ко-монотонность кусочно-линейных функций) при любом изменении последовательности координатных осей при изображении объекта в системе параллельных координат [183].

Процедуру кластеризации, результат которой не зависит от последовательности координатных осей показателей исследуемых объектов, будем называть «порядково-инвариантной паттерн-кластеризацией», а кластеры, полученные в результате ее применения, соответственно, порядко-инвариантными паттерн-кластерами. Основные свойства данной процедуры представлены в [184], а методология корректировки результатов при наличии погрешности в исходных данных в [185] Представленная ниже теорема формулирует важное свойство порядково-инвариантных паттерн кластеров.

Теорема 2.1. Порядково-инвариантные паттерн-кластеры не пересекаются. Иными словами, одна строка данных не может принадлежать двум различным порядково-инвариантным паттерн-кластерам.

Теорема фактически определяет однозначность порядково-инвариантной паттерн-кластеризации, а потому имеет важное приложение как для построения алгоритма ее реализации, так и для разработки отдельных приложений.

2.3.3 Примеры порядково-инвариантных паттерн-кластеров с взаимно симметричной структурой кусочно-линейных функций

Цель данного пункта – продемонстрировать (на примере широко известных, практически, “академических” наборах данных) существование кластеров с симметричной структурой паттернов – с одной стороны, и отдельные стороны предлагаемого метода для целей Visual Data Mining – с другой. Для этих целей использованы три различных наборы данных (основные результаты представлены в [186]).

1. «Balance Scale Data Set» [187].

Таблица исходных данных содержит 625 строк. Каждая строка отражает значения пяти показателей (p_0, p_1, p_2, p_3, p_4), характеризующих параметры и состояние рычажных весов.

Параметр p_0 = {«*B*», «*L*», «*R*»} характеризует следующие состояния: равновесие; отклонение влево или вправо.

Оставшиеся параметры p_1, p_2, p_3 и p_4 принимают целочисленные значения от одного до пяти и характеризуют:

– p_1={1, 2, 3, 4, 5} – число весовых единиц левой части;

– p_2={1, 2, 3, 4, 5} – расстояние от точки опоры слева;

– p_3={1, 2, 3, 4, 5} – число весовых единиц правой части; и

– p_4={1, 2, 3, 4, 5} – расстояние от точки опоры справа.

Набор исходных данных содержит 49 строк, соответствующих состоянию «B» (равновесие); 288 – состоянию «L» (отклонение влево); и 288 – состоянию «R» (отклонение вправо).

Цель эксперимента – используя набор данных «Balance Scale Data Set» продемонстрировать отдельные возможности метода порядково-инвариантной паттерн-кластеризации для визуального анализа данных при решении задачи Data Mining, выявить кластеры, обладающие свойством симметрии паттернов данных.

Результатом работы процедуры порядково-инвариантной кластеризации явилось разделение исходного набора данных «Balance Scale Data Set» на множество порядково-инвариантных паттерн-кластеров, обладающих рядом визуально отличительных особенностей, позволяющих подразделить их на три группы [188].

Прежде всего, выделяется группа порядково-инвариантных паттерн-кластеров, для которых характерна абсолютная симметрия правой и левой части графиков (см. рисунок 2.10 ниже). Суммарно, эта группа объединяет 45 из 49 строк, соответствующих состоянию «B» (равновесие).

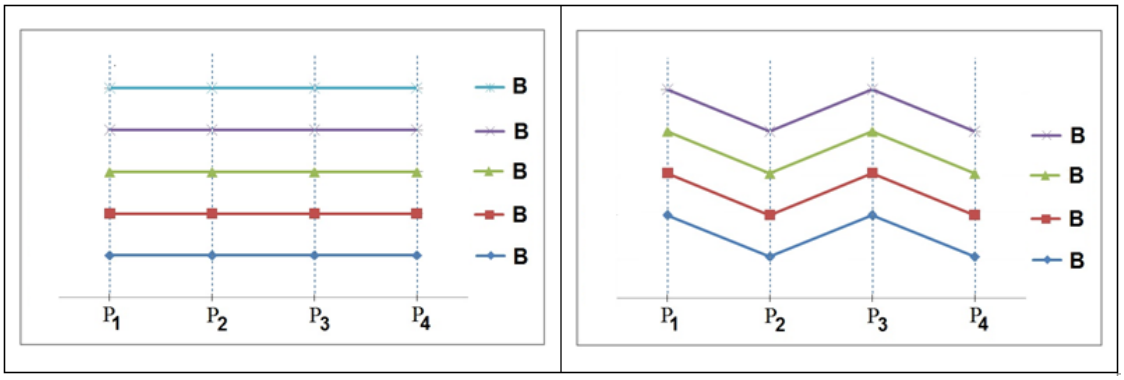


Рисунок 2.10 – Примеры объектов порядково-инвариантных паттерн-кластеров, соответствующих состоянию равновесие

Оставшееся множество порядково-инвариантных паттерн-кластеров визуально подразделяется на четыре группы, обладающие следующей ярко выраженной особенностью: каждому порядково-инвариантному паттер-кластеру одной группы можно указать соответствующие ему порядково-инвариантные паттерн-кластеры трех других групп, обладающих зеркально симметричной формой кусочно-линейных функций относительно вертикальных и горизонтальных осей (см. рисунок 2.11).

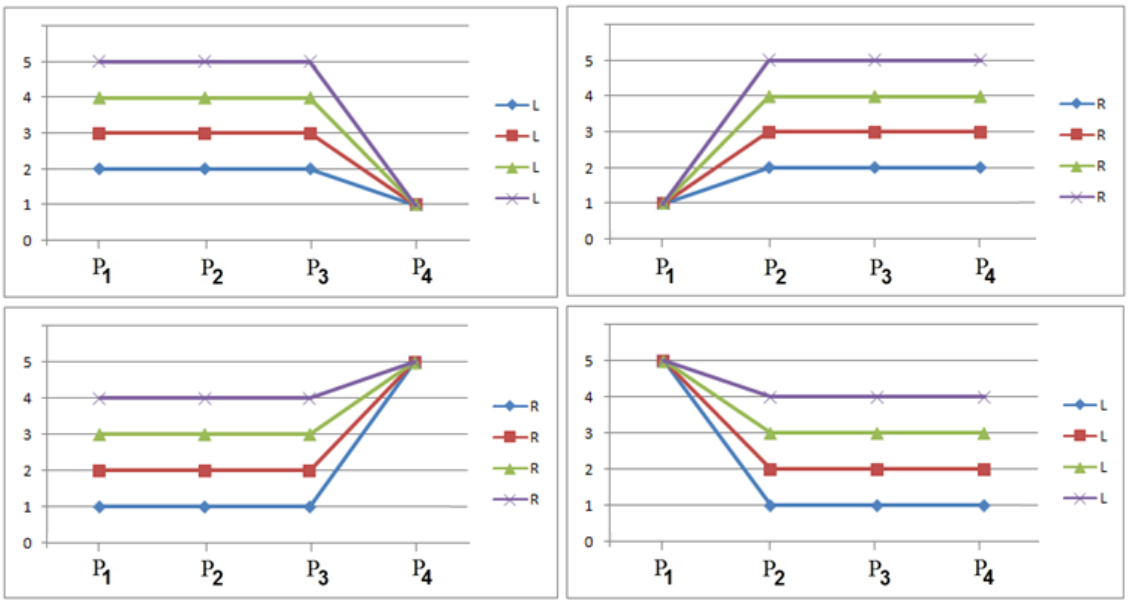


Рисунок 2.11 – Четыре порядково-инвариантных паттерн кластера, характеризующихся зеркально симметричной формой графиков

Важно отметить, что практически все порядково-инвариантные паттерн-кластеры содержали объекты только одного типа: «B», «L» или «R». Исключение составили лишь четыре объекта группы «B» (рисунок 2.12). В отличие от других объектов этой группы, они были причислены к порядково-инвариантным паттерн-кластерам, содержащих объекты с зеркально симметричной формой кусочго-линейных функций.

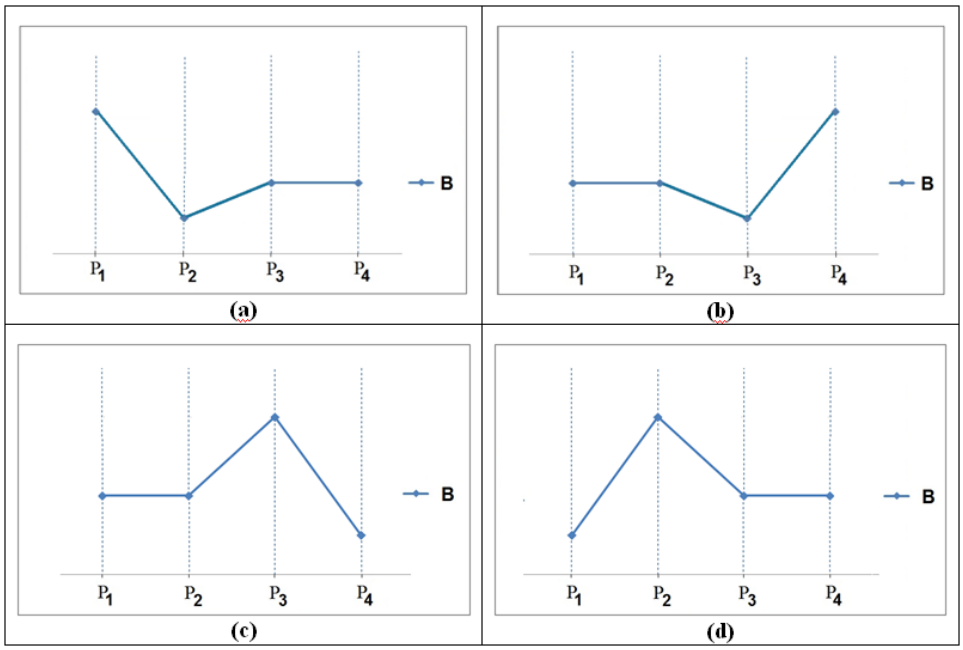


Рисунок 2.12 – Объекты группы «B» с зеркально симметричной формой кусочно-линейных функций

Таким образом, применение порядково-инвариантной паттерн-кластеризации к набору данных Balance Scale Data Set позволило разделить исходные данные на три группы, соответствующие различным состояниям: равновесие, отклонение влево и отклонение вправо. Вид кусочно-линейных функций позволяет сделать предположение о наличии определенных симметричных свойств исходного объекта.

2. «Car Data Set».

Набор данных «Car Data Set» содержит технические характеристики автомобилей, производимых различными фирмами США, Европы и Японии в период 1970 – 1982. Впервые представлен на American Statistical Association Data Exposition в 1983 году.

Описание исходного набора данных. Таблица исходных данных содержит 406 строк. Каждая строка отражает значения девяти параметров (p_0, p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8), характеризующие следующие параметры автомобиля:

Параметр p_0 отражает модель автомобиля. Параметры p_1, p_2, p_3, p_4, p_5 и p_6 принимают числовые значения и описывают технические характеристики автомобиля:

– p_1 – расход топлива;

– p_2 – число цилиндров;

– p_3 – объем двигателя;

– p_4 – мощность;

– p_5 – вес автомобиля;

– p_6 – время разгона от 0 до 60 миль в час.

Параметры p_7 и p_8 характеризуют:

– p_7 – год выпуска модели;

– p_8 – происхождение: США, Европа или Япония.

Предпосылки экспериментального исследования. Визуальный анализ больших объемов информации и выявление в ней отдельных групп, обладающих схожими чертами, существенно затруднен вследствие наличия шума

в изображении при отображении значительного числа данных. В качестве примера, приведем графическое представление набора данных «Car Data Set» (рисунок 2.13).

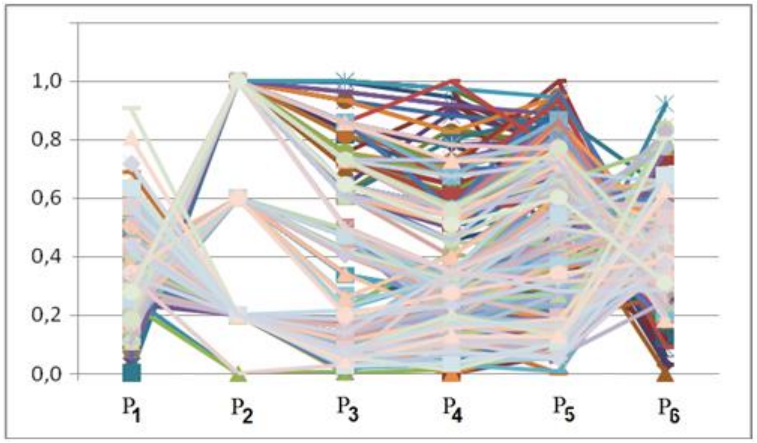


Рисунок 2.13 – Графическое представление набора нормализованных данных «Car Data Set»

Снижению шума изображения с целью облегчения визуального анализа на примере набора данных «Car Data Set», посвящена, в частности, работа [189]. Однако, снижение шума не решает полностью проблему, поскольку само наличие шума связано, прежде всего, с отображением и взаимным наложением больших объемов разнотипных графических данных.

Цель эксперимента. Используя набор данных «Car Data Set», продемонстрировать возможности рассматриваемого метода для снижения шума визуализации данных и выявления взаимно симметричных структур.

Результаты эксперимента. Кластеризация проводилась по техническим характеристикам автомобилей p_1, p_2, p_3, p_4, p_5 и p_6, выраженных в нормализованных величинах в диапазоне [0,1]:

$$p_{in} = \frac{p_i - p_{i\min}}{p_{i\max} - p_{i\min}},$$

где $p_{i\min} = \min_j(p_{ij}), \quad p_{i\max} = \max_j(p_{ij}) \quad | \; i = 1,...,6; \quad j = 1,...,406$.

Результатом явилось разбиение исходного множества на ряд кластеров, примеры которых показаны на рисунке 2.14 ниже.

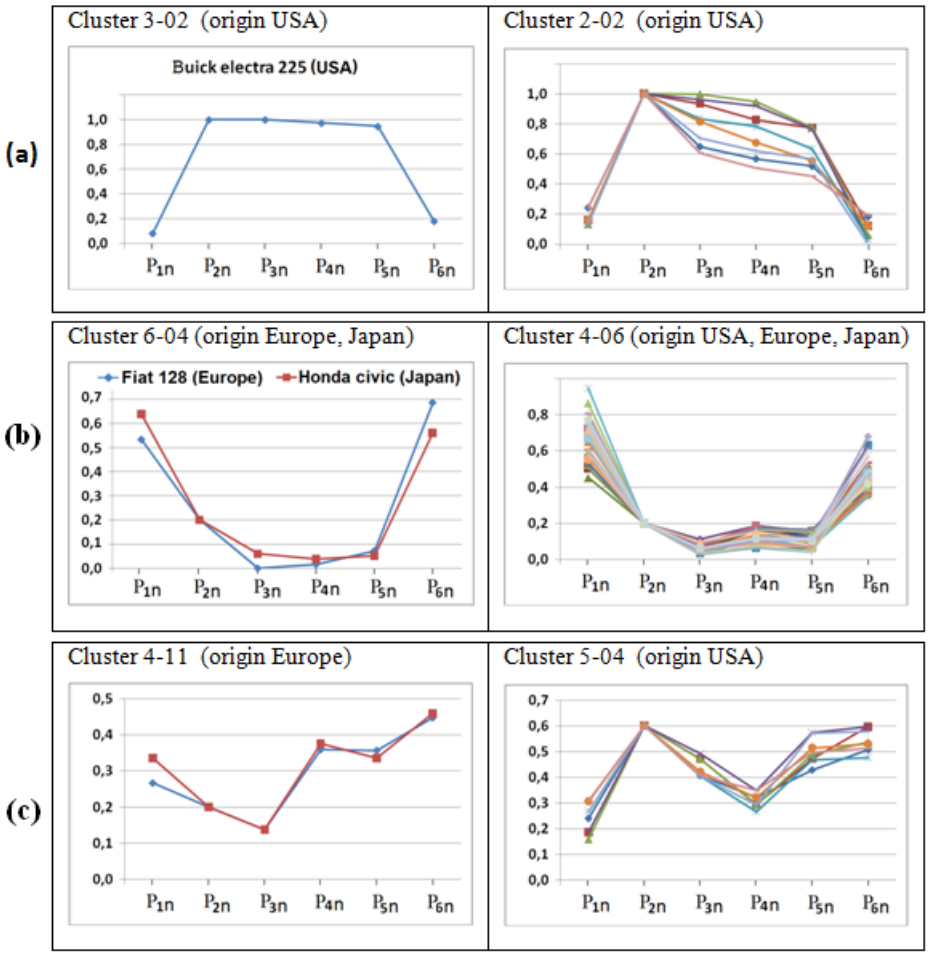


Рисунок 2.14 – Визуально различимые виды паттернов: (a) выпуклые вверх и их вариации; (b) выпуклые вниз и их вариации; (c) «промежуточные» виды

Приведенные рисунки демонстрируют существенное снижение шума изображения, что является следствием двух факторов:

– разделение исходного набора на ряд порядко-инвариантных паттерн-кластеров, содержащих структурно близкие данные;

– относительно небольшим (по сравнению с исходным набором) числом данных в каждом порядково-инвариантном паттерн-кластере.

Дальнейший анализ показал, что выделенные порядково-инвариантные паттерн-кластеры представляют две группы данных:

– первая группа представляла порядково-инвариантные паттерн-кластеры, содержащие характеристики автомобилей, производимых только в США;

– вторая группа представляла порядково-инвариантных паттерн-кластеры, содержащие характеристики автомобилей, производимых в Европе, Японии и США.

Оценивая возможность использования результатов порядково-инвариантной паттерн-кластеризации для целей Visual Data Mining, следует отметить, что визуально выделяются следующие типовые виды паттернов (см. рисунок 2.14):

– паттерны выделяемых кластеров имеют вид кусочно-линейных функций, выпуклых вверх. Эти кластеры характерны для автомобилей США начала 70 годов. Модификация отдельных параметров автомобилей, производимых в последующие годы, отражалась на видоизменении паттерна, что продемонстрировано на рисунке 2.14a;

– другой тип визуально выделяемого паттерна имеет характер прямо противоположный предыдущему и представлен кусочно-линейной функцией, выпуклой вниз. Выделяемые порядково-инвариантные паттерн-кластеры характерны для ряда моделей автомобилей, выпускаемых в Европе, Японии и США начиная со второй половины 70 годов. Как и в предыдущем случае, модификация отдельных параметров автомобилей, производимых в последующие годы отражалась на видоизменении паттерна, что продемонстрировано на рисунке 2.14b;

– третий вид паттерна является “промежуточным” между рассмотренными выше. Для него характерны как порядково-инвариантные паттерн-кластеры, содержащие данные автомобилей, выпускаемых только в США, так и порядково-инвариантные паттерн-кластеры, в которых представлены автомобили различных стран (рисунок 2.14c).

Наблюдаемая симметрия структуры данных отражает различные подходы в создании отдельных моделей автомобилей в США, Европе и Японии того времени. В США, обладающих относительно высокими запасами энергетических ресурсов, активно выпускались мощные и тяжелые модели автомо-

билей, требующие использования 8-ми (в некоторых моделях – 6-ти) цилиндровых двигателей, высокого расхода топлива, и обладающих невысоким ускорением при разгоне. Для таких моделей характерны кусочно-линейные функции, выпуклые вверх (рисунок 2.14a). В Европе и Японии предпочтение отдавалось более легким и «юрким» моделям автомобилей, с использованием 4-х (в некоторых моделях, даже, 3-х) цилиндровых двигателей, не требующим столь высокого расхода топлива. Для них характерны кусочно-линейные функции, выпуклые вниз (рисунок 2.14b). Развитие автомобилестроения приводило к сближению этих различных подходов, что и отразилось на виде паттернов (рисунок 2.14c).

3. «Pollen Data Set».

Синтетический набор данных Pollen Data Set создан David Coleman at RCA Laboratories in Princeton, штата New Jersey.

Описание исходного набора данных. Таблица исходных данных содержит 3848 строк. Каждая строка отражает значения шести параметров (p_1, p_2, p_3, p_4, p_5, p_6), последний из которых – p_6 носит дополнительный характер, отражающий номер строки в наборе данных. Согласно описанию переменных p_1 – p_5, данному автором, «the variables were given entirely ficticious names» и они условно характеризуют геометрические размеры вдоль осей x, y, z, массу и плотность зерен пыльцы:

– p_1 – ridge;

– p_2 – nub;

– p_3 – crack;

– p_4 – weight;

– p_5 – density;

– p_6 – observation number (for convenience).

Цель эксперимента – используя набор данных «Pollen Data Set» продемонстрировать возможности рассматриваемого метода для относительно больших объемов данных.

Предпосылки экспериментального исследования. Набор данных «Pollen Data Set» является синтетическим, т.е. относится к классу искусственно формируемых наборов, для которых трудно ожидать наличия кластеров с взаимно симметричной структурой данных. Отметим, также, сложность визуального анализа данных, что связано с большим объемом и высокой плотностью (см. рисунок 2.15a). Использование в изображении градаций интенсивности серого цвета, предложенное в работе [190] позволило частично решить указанную проблему и улучшить визуальное представление (см. рисунок 2.15b).

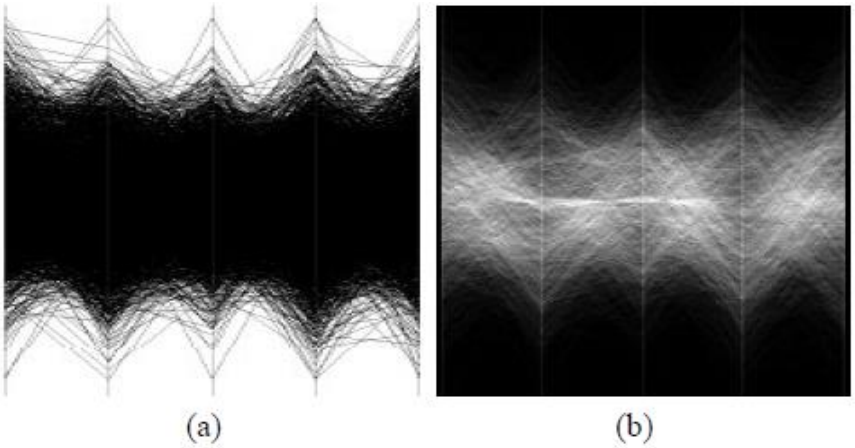


Рисунок 2.15 – a) визуализация исходного набора данных; b) визуализация исходного набора данных с различной интенсивностью уровня серого (рисунок взят из [190])

Дальнейшее развитие метод получил в работе [191], где авторами предложены интерактивные алгоритмы Interactive Parallel Coordinates Frequency Plot и Interactive Parallel Coordinates Density Plot, в которых уровни интенсивности серого определяются на основе анализа частоты и плотности участков изображения. Пример применения алгоритма Interactive Parallel Coordinates Density Plot к набору данных Pollen Data Set показан на рисунке 2.16.

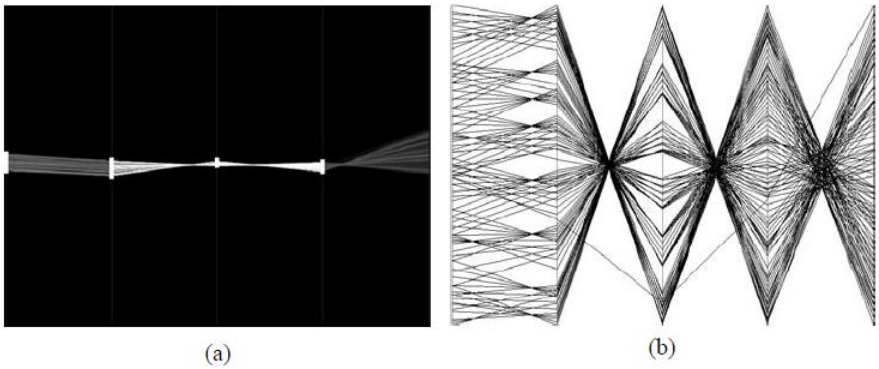


Рисунок2.16 – a) визуализация данных с IPC Density Plot; b) увеличение рисунка (взято из [190])

Визуальный анализ рисунков 2.15b и 2.15b, позволяет сделать предположение о наличии в наборе Pollen Data Set групп объектов с симметричной структурой паттернов данных.

Результаты проведенного эксперимента подтвердили наличие в рассматриваемом наборе кластеров с взаимно симметричной структурой паттернов данных, что отражено на рисунке 2.17.

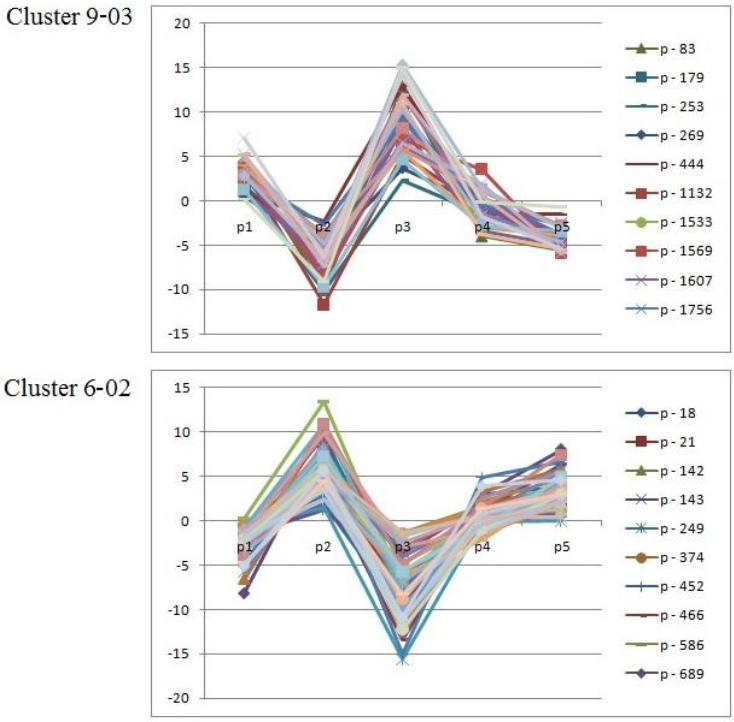


Рисунок 2.17 – Примеры взаимной симметрии структуры кусочно-линейных функций порядково-инвариантных паттерн-кластеров в наборе данных Pollen Data Set

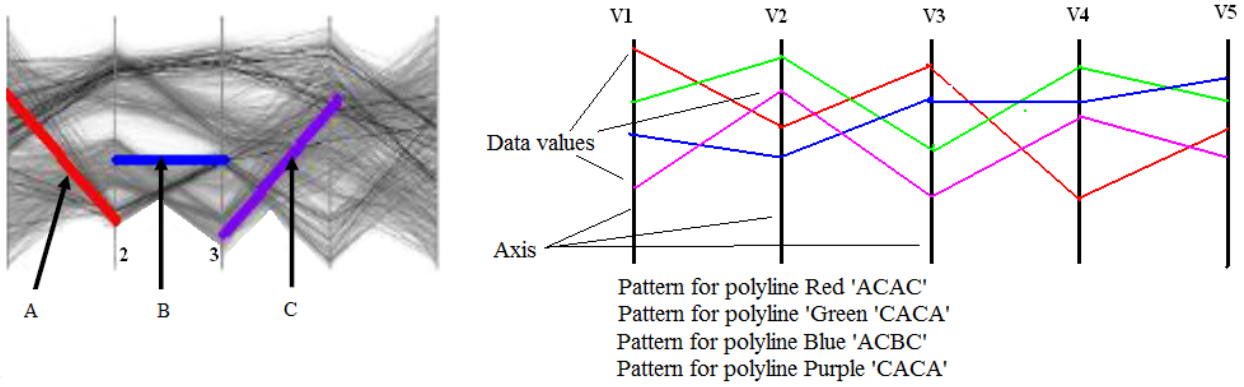
Наличие таких структур с достаточно большим числом элементов, обусловлено, использованным методом формирования набора данных Pollen Data

Set. Согласно описанию, один из этапов состоял в том, что после формирования части этого набора (части A), вторая его часть (часть B) была получена путем инвертирования знака у трех из пяти переменных части A, что фактически, означало зеркальное отражение данного параметра относительно нулевой точки соответствующей оси.

2.3.4 Выделение порядково-инвариантных паттерн-кластеров с взаимно симметричной структурой данных.

Выделение порядково-инвариантных паттерн-кластеров с взаимно симметричной структурой кусочно-линейных функций реализуется в два этапа. На первом этапе используется порядково-инвариантная паттерн-кластеризация к исходному набору данных. Каждый порядково-инвариантный паттерн-кластер, в этом случае, содержит данные, графически представляемые в системе параллельных координат ко-монотонными графиками. В силу порядковой инвариантности получаемых групп объектов, ко-монотонность сохраняется для любой последовательности координатных осей.

На втором этапе, осуществляется поиск кластеров, обладающих свойством взаимной симметрии структуры кусочно-линейных функций. Для небольшого числа сформированных кластеров поиск может быть выполнен визуально. Учтем, что каждый формируемый на первом этапе кластер содержит набор данных, графически представляемых в системе параллельных координат в виде ко-монотонных кусочно-линейных функций. Общий вид данных кусочно-линейных функций каждого кластера может быть описан кодовой последовательностью, каждый символ которой («A», «B» или «C») характеризует тип наклона линии на отдельном интервале [182], см. рисунок 2.18.



(a) (b)

Рисунок 2.18 – (a) категория наклона A, B и C в системе параллельных координат; (b) кусочно-линейные функции, характеризующие различные паттерны (рисунок взят из [182])

Для поиска кластера содержащего инверсную (относительно горизонтальной оси) форму кусочно-линейной функции, следует сформировать характеризующую его «инверсную» кодовую последовательность символов, заменяя символы «A» на «C» и «C» на «A» (оставляя символ «B» без изменения).

Для формирования кодовой последовательности кластера характеризующегося кусочно-линейной функции, симметричной относительно вертикальной оси, следует записать инверсную кодовую последовательность символов в обратном порядке.

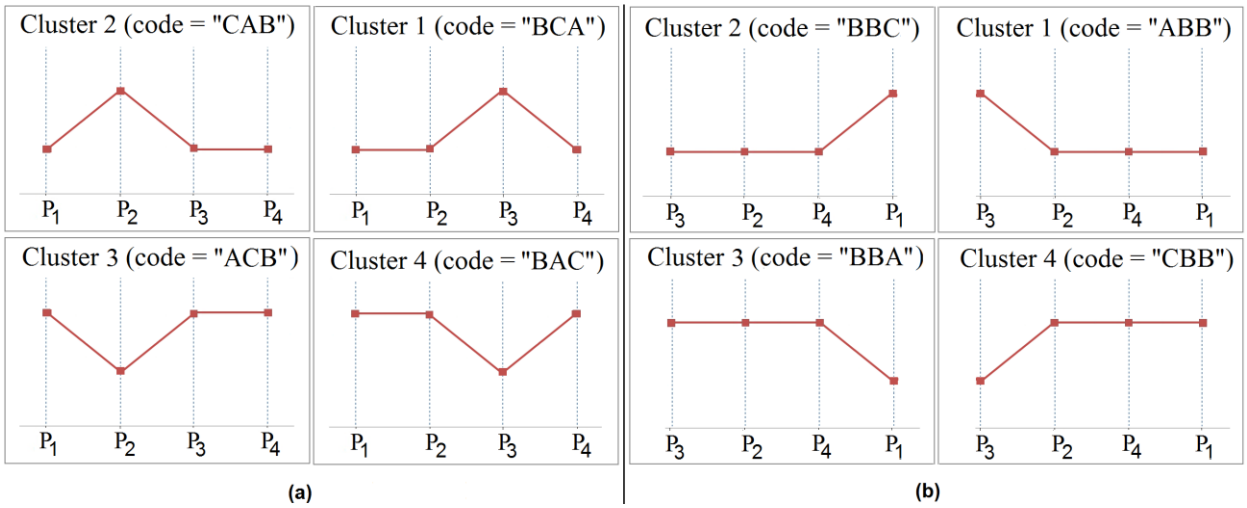


Рисунок 2.19 – (a) четыре порядко-инвариантных паттерн-кластера с взаимно симметричной структурой; (b) изменение порядка координатных осей меняет вид кусочно-линейных функций, но не меняет их симметричность

На рисунке 2.19a представлены четыре порядково-инвариантных паттерн-кластера с взаимно симметричной структурой. Изменение порядка координатных осей меняет форму кусочно-линейных функций, но не нарушает их симметрию, см. рисунок 2.19b.

2.3.5 Некоторые выводы

В работе рассмотрены вопросы поиска и выделения групп объектов, обладающих свойством взаимной симметрии структуры кусочно-линейных функций. Предложенный алгоритм позволяет выделять зеркальные или визуально симметричные структуры в условиях относительно большого объема данных и высокого шума изображения. При этом симметрия паттернов сохраняется при любом порядке координатных осей.

Приведены результаты применения описываемого метода на примере наборов данных Balance Scale Data Set, Car Data Set и Pollen Data Set. Отмечается, что в каждом рассмотренном примере, симметрия паттернов имела определенное смысловое содержание.

Рассмотренные примеры демонстрируют потенциальную возможность извлечения дополнительной информации об объектах на основании анализа структуры порождаемых ими данных. Отметим, также, что визуально выделяемые отличительные признаки близки к тем, которыми оперирует человек. Эта особенность позволяет рассматривать возможность использования описываемого метода, например, в системах моделирования «человеческого» выбора.

3 Исследования организационных моделей и механизмов

В этом разделе отчета приведены результаты по исследованию организационных моделей и механизмов. Полученные результаты отражены в следующих подразделах:

– исследования конкурентной кооперации в Арктике;

– исследование моделей по оптимальному размещению объектов и привлечению частных инвестиций.

3.1 Исследования конкурентной кооперации в Арктике

Проект «Общество и передовые технологии в Арктике» (SATA) разработан для обучения норвежских и российских аспирантов и студентов интернациональному междисциплинарному анализу и решению проблем, касающихся общества и передовых технологий в Арктике. Задача – междисциплинарное применение фундаментальных исследований для улучшения условий жизни в Арктике с использованием передовых технологий (особенно дистанционного зондирования и космических технологий). Перед аспирантами стоит задача разработать социально-технические решения, ориентированные на потребности и возможности общества, экономики, культуры, регулирования, естествознания и техники.

Полевой курс 2018 года проходил в Тромсё, на базе UiT-Арктического университета Норвегии. В течение недели интенсивных исследований и взаимодействия с ведущими исследовательскими центрами в различных областях научно-исследовательской деятельности для арктических регионов студенты, в формате соревнования между проектными группами, разработали несколько исследовательских предложений. Проведенный в течение этого года анализ сформированных предложений показал, что они оказались тесно взаимосвязанными, образуя круговую структуру, которую можно назвать по имени всего

курса «цикл SATA» – взаимодействие между промышленностью, наукой, образованием, обществом, являющейся наглядной иллюстрацией феномена конкурентного сотрудничества (coopetition) – см. [192, 193].

Ниже мы представляем краткий обзор предложений и даем краткое объяснение самой идеи «цикла».

Проект 1 «Ранжирование мест размещения нефтяных платформ в Баренцевом море: оценка полезности для заинтересованных сторон», разработанный Сергеем Деминым, Кэт Ходжсон, Яной Коротовой и Денисом Тверским. Мы рассматриваем это как «Технологическое влияние» (PTe).

Целью данного проекта является создание системы ранжирования, которая оценивает несколько платформ для потенциальной добычи нефти на основе их полезности. С этой целью команда проекта постарается максимально использовать потенциальные возможности и минимизировать потенциальные риски, а также конфликты между различными заинтересованными сторонами.

Проект 2 «Сеть арктических исследовательских центров», разработанный Яной Зыковой, Кёнг Ри, Гузель Альмухаметовой, Марией Ермоловой. Мы рассматриваем это как «научное влияние» (PSc).

Целью данного проекта является повышение уровня сотрудничества научных центров, занимающихся проблемами Арктики. В рамках исследования предполагается:

– определить уровень сотрудничества между исследовательскими центрами, работающими по одним и тем же темам, касающимся Арктического региона, используя разные критерии, например, количество совместных публикаций, количество совместных проектов, программы обмена студентами в рамках каждой дисциплины и трансдисциплинарности;

– определить основные направления исследований исследовательских центров; какие темы менее освещены и нуждаются в большем международном сотрудничестве;
– определить факторы и их влияние на уровень сотрудничества, в том числе экономические, научные, культурные и географические факторы;
– проанализировать спрос на исследования в различных областях на основе проблем и потребностей страны (изменения климата, уровень здоровья, другие проблемы Арктики).

Проект 3 «Циркуляция знаний в Арктике: места учебы, проживания и занятий северян», разработанный Даниэлой Шустер, Юлией Гуровой, Ириной Гавриленковой, Александром Проскуряковым. Мы рассматриваем это как «образовательное влияние» (PEd).

В рамках проекта предполагается междисциплинарное исследование карьерных траекторий выпускников вузов. Основная цель – рассмотреть вопрос о перемещении выпускников между местами их рождения, учебы и работы и определить причины их перемещения. Моделирование этих процессов позволит нам отслеживать людские потоки и прогнозировать будущие варианты учебы, работы и проживания молодых людей в Арктике. Этим можно разрешить многие национальные и международные конфликты между пользователями в районе Баренцева моря, предлагая стратегии привлечения квалифицированной рабочей силы, что требует интенсивного сотрудничества различных заинтересованных сторон.

Проект 4 «Миграция в Арктике», разработанный Александром Бульдяевым, Хема Надераджа, Анной Резяповой, Юань Цзин, Йоханом Биркелундом. Мы рассматриваем это как «социальное влияние» (PSo).

Целью этого проекта является продвижение и распространение знаний о причинах внутренней миграции в норвежских и российских арктических регионах. Общество в Арктике сталкивается с различными последствиями процесса развития, такими как изменение климата, увеличение разнообразия в обществе, рост новых технологий и нехватка природных ресурсов. Проект посвящен изучению того, как миграция в Арктике будет способствовать решению проблем, упомянутых выше.

“Цикл SATA”

Междисциплинарность курса SATA, сильная взаимосвязь всех дисциплин, которые были предоставлены студентам в одно и то же время выявили существование очевидных взаимосвязей между различными проектами, которые позволяют сформировать круговое представление, названное нами циклом SATA по названию всего курса.

Проект 1 «Ранжирование мест размещения нефтяных вышек в Баренцевом море: оценка полезности для многих заинтересованных сторон» и проект 2 «Сеть арктических исследовательских центров» образуют научно-технологическую связь (LTeSc). Проблемы, вызванные арктическими особенностями промышленного актива, вызывают спрос на специальные исследовательские темы, а соответствующая исследовательская деятельность вносит большой вклад в возможность решения промышленных задач.

Проект 2 «Сеть исследовательских центров Арктики» и проект 3 «Циркуляция знаний в Арктике: места учебы, проживания и занятий северян» образуют научно-образовательную связь (LScEd). Возможность достижения самых современных научных результатов в значительной степени зависит от наличия у исследователей соответствующих навыков и опыта, что, в свою очередь, в значительной степени определяет потоки выпускников вузов и их карьерные траектории. В свою очередь, запуск амбициозных исследовательских проектов позволяет создавать новые образовательные программы, которые

могут быть весьма конкурентоспособными, чтобы привлечь больше талантливых студентов в арктические университеты.

Проект 3 «Циркуляция знаний в Арктике: места учебы, проживания и занятий северян» и проект 4 «Миграция в Арктике» образуют связь Образование-Общество (LEdSo). Очевидно, что циркуляция «мозгов» способствует полной миграции, но в то же время важно учитывать, как глобальная миграция влияет на миграционные потоки студентов и научных исследователей. Наконец, проект 4 «Миграция в Арктике» и проект 1 «Ранжирование мест размещения нефтяных платформ в Баренцевом море: оценка полезности для многих заинтересованных сторон» образуют связь Общество-Технология (LSoTe). Влияние на окружающую среду технологий, применяемых в промышленной деятельности, влияет на качество жизни, что, в свою очередь, является одним из фундаментальных факторов, определяющих миграционные потоки. В свою очередь, высокие стандарты качества жизни, которые позволяют иммиграции превалировать над эмиграцией, стимулируют высокие стандарты безопасности и воздействия на окружающую среду промышленной деятельности, требующей внедрения более совершенных технологий.

Рисунок 3.1 – Цикл SATA

Возможная иллюстрация цикла SATA представлена на рисунке 3.1. Очевидно, что структура этого замкнутого графа не является жесткой в смысле узлов и ребер и может и должна быть расширена в ходе будущей деятельности

SATA. Но даже в этот момент это уже иллюстрирует ожидаемую междисциплинарную и тесную взаимосвязь студенческих проектов, реализуемых в рамках проекта SATA.

3.2 Исследование моделей по оптимальному размещению объектов и привлечению частных инвестиций

3.2.1 Введение

Как известно, одна из основных функций системного анализа состоит а) в отыскании сходства в системах различной природы, б) в изучении классов схожих проблем и в) в описании особенностей, приписываемых всем проблемам из одного и тот же класса [194, 195]. Действительно, из общих соображений ясно, что, например, оптимизация размещения объектов самой разной природы должна осуществляться в соответствии с некоторыми общими принципами, не зависящими от природы объектов, которые (принципы) могут модифицироваться с целью отразить в выбираемом размещении именно эту природу. Хотя, разумеется, желательно отыскивать возможно более широкие классы систем, обладающих определенным сходством, выделение даже двух таких схожих систем, представляющих, однако, интерес для практики, заслуживает внимания. Транспортные и энергетические системы дают пример таких схожих систем, для которых сходства в функционировании достаточно очевидны и позволяют изучать это функционирование на основе одних и тех же математических моделей. Подобная общность принципов очевидна и в привлечении инвестиций в системы самой различной природы, и транспортные и энергетические системы, как примеры крупномасштабных систем, снова допускают исследование их инвестиционных проблем на основе одних и тех же математических моделей.

Например, такая "чисто энергетическая" проблема как проблема передачи электроэнергии от нескольких электростанций к нескольким потребите-

лям может быть математически сформулирована как однопродуктовая транспортная задача. Такая "чисто энергетическая" система как станция замены электрических батарей, обслуживающая нескольких пунктов а) из каждого из которых истощенные батареи разного химического состава, емкостей и размеров забираются и доставляются на эту станцию, где эти батареи перезаряжаются, и б) к каждому из которых перезаряженные на станции батареи перевозятся, может рассматриваться как транспортный узел обработки грузов. Здесь проблемы взаимодействия разных типов батарей с различными видами технологического оборудования, развернутого для обслуживания процессов зарядки, похожи на проблемы взаимодействия различных видов грузового транспорта в транспортных узлах. Наконец, различные расчеты для электрических сетей, особенно для тех, которые используют как традиционные, так и возобновляемые источники энергии, оказываются аналогичными соответствующим расчетам для грузовых транспортных сетей.

Хотя, конечно, энергия и транспорт имеют проблемы, характерные только для одного из этих двух видов систем или для другого, представляется разумным изучить проблемы, которые могут возникнуть в обеих этих системах. Это представляется разумным, даже если в настоящее время, по крайней мере, некоторые проблемы являются предметом отдельного интенсивного исследования лишь в каждой из этих двух систем или вообще рассматриваются только в одной из них, хотя они уже начали привлекать внимание исследователей, изучающих другую систему.

Проблемы, связанные с созданием и развитием инфраструктуры в энергетике и транспортных системах образуют класс таких проблем. В частности, проблемы а) распределения транспортных узлов в регионе, б) выбора пропускных способностей (мощностей) этих транспортных узлов и схем перемещения грузов через эти узлы, в) оценки общего объема расходов, ожидаемых в связи со строительными и ремонтными работами в этих узлах, и г) оценки ожидаемого дохода, который функционирующие транспортные узлы могут генериро-

вать, являются типичными для транспортных систем. Эти проблемы представляют интерес как для государственных администраций, так и для частных инвесторов. Таким образом, можно ожидать, что создание и развитие инструментария, позволяющего делать вычисления, связанные со всеми вышеупомянутыми мероприятиями, поможет каждой государственной администрации в ее переговорах с частными инвесторами об их участии в крупномасштабных проектах, важных для общества.

Практически те же проблемы должны рассматриваться в отношении станций замены батарей. Именно, а) размещение станций замены батарей в определенном регионе или муниципалитете, б) выбор мощностей этих станций замены и схем перемещения истощенных и перезаряженных батарей через них, в) оценка общего объема расходов, которые, как ожидается, потребуются для строительства и обеспечения финансирования деятельности по обслуживанию станций замены батарей, и г) оценка ожидаемого объема дохода, который функционирующие станции замены батарей могут генерировать, являются типичными для энергетических систем.

Все рассматриваемые настоящем исследовании проблемы, общие для транспортных и энергетических систем, сформулированы либо в «транспортных терминах», либо в «энергетических терминах», в зависимости от того, какая именно терминология представляется авторам исследования более естественной. Тем не менее, проблемы, сформулированные в «транспортных терминах» можно легко переформулировать в «энергетических терминах», и наоборот, проблемы, сформулированные в «энергетических терминах» можно переформулировать в «транспортных терминах», в которых все полученные результаты могут быть интерпретированы.

В развитии региональных инфраструктур грузовых перевозок, создание комплекса новых транспортных узлов с подъездными путями к ним является одной из двух ключевых частей инженерного проекта, который региональные администрации могут предложить финансировать своим потенциальным

партнерам из частного сектора. Другая ключевая часть проекта связана с обеспечением эффективного управления развитыми таким образом транспортными инфраструктурами. Только эти две части проекта развития региональной инфраструктуры грузоперевозок—обеспечение комплекса строительных и инженерных работ, связанных со строительством транспортных узлов и подъездных путей к ним, и эффективного управления этими элементами региональной инфраструктуры грузовых перевозок— являются предметом рассмотрения в настоящем исследовании.

Хорошо известно, что прикладная математика и компьютерные автоматизированные системы помогают в анализе и решении логистических, эксплуатационных и управленческих проблем в транспортных системах. Тем не менее, инвестиционные проблемы, связанные с логистикой, лежащей в основе региональной грузовой транспортной инфраструктуры, имеют свою специфику и особенности. Эти особенности усложняют использование общего финансового инженеринга к выработке подходов к решению проблем стратегического управления в транспортных (и энергетических) системах, с которыми сталкивается каждая государственная администрация, в частности, проблем оценки требуемого объема инвестиций в транспортную (и энергетическую) инфраструктуру муниципального образования. Вопрос ставится следующим образом: Можно ли создать инструментарий для поддержки принятия решений, который базируясь на общих принципах финансового анализа, будет отражать упомянутые особенности и специфику грузовой транспортной (энергетической) инфраструктуры при оценке объема инвестиций, необходимых региональным администрациям, как от государственного, так и от частного сектора, для эффективного развития и управления конкретной региональной транспортной (энергетической) инфраструктурой?

Одной из целей настоящего исследования является попытка продемонстрировать, что в отношении региональных инфраструктур грузовых перевозок ответ положительный, при условии, что соответствующие аналитические

средства правильно выбраны и правильно применяются. В частности, авторами предложена математическая модель, лежащая в основе инструментария поддержки принятия решений для оценки необходимого объема инвестиций в развитие региональной инфраструктуры грузоперевозок, позволяющего региональной администрации начать переговоры с потенциальными инвесторами из частного сектора по финансированию соответствующего проекта. Предложенная модель является нелинейным обобщением известной модели, используемой при изучении проблем размещения объектов, и отражает основные правовые, инженерные и финансовые возможности региональной администрации по организации сотрудничества с частным сектором в рамках, например, потенциального государственно-частного партнерства. На основе предложенной (обобщенной) модели, оценка расходов, связанных с реализацией какого-либо конкретного проекта, может быть получена с использованием стандартного программного обеспечения для решения задач оптимизации. При этом решения для соответствующей задачи оптимизации могут быть достаточно быстро получены, когда часть данных (или все данные), отражающих (отражающие) географию соответствующего региона, может (могут) быть известна (известны) только приблизительно (что отражает условия неопределенности, при которых вышеуказанные расходы оцениваются). В настоящем исследовании демонстрируются результаты применения предложенной модели и стандартного программного обеспечения на модельных примерах грузовых транспортных систем и формулируются некоторые рекомендации по их использованию. Результаты, представленные в настоящем отчете, получены в рамках исследований, являющихся продолжением исследований, проведенных автором в 2017-2018 гг. (при участии Г.Г. Федина). В частности, в нем приводится модифицированная математическая модель, впервые упомянутая в отчете автора 2017 года, и обсуждаются результаты экспериментальных расчетов, проведенных с помощью этой модели в 2018-2019 гг. и опубликованных в [194].

Другой целью настоящего исследования явилось выявление и краткое описание классов системных проблем, связанных с формированием государственно-частного партнерства (на примере энергетических систем), опубликованное в [195]. Исследование математических проблем, связанных с получением закономерностей и количественных оценок при анализе и решении прикладных проблем в рамках каждого из классов, может представить предмет дальнейших исследований в области механизмов формирования государственно-частного партнерства.

Наконец, в рамках настоящего исследования были модифицированы математические модели, на основе которых переформулированы математические задачи оценки объема инвестиций, требуемого для размещения цепи зарядных станций электромобилей на шоссе, впервые содержательно сформулированные в отчете 2017 года, и доказаны соответствующие математические утверждения, опубликованные в статье [196].

3.2.2 Разработка и исследование модифицированной модели размещения грузового транспортного узла

Как правило, география региона и уже существующая транспортная инфраструктура страны определяют потенциальные места, в которых новые транспортные узлы могут быть построены. Когда это так, ожидаемые объемы, например, ежегодных грузопотоков через эти новые транспортные узлы помогают примерно оценить желаемые мощности новых узлов. Тем не менее, пропускные способности, как новых транспортных узлов, так и подъездных путей к ним влияют на распределение ожидаемого общего грузопотока в новой грузовой транспортной инфраструктуре в целом, которую планируется разработать.

Таким образом, при рассмотрении вопроса о создании и развитии новой региональной инфраструктуры перевозок грузов, инструмент поддержки принятия решений для анализа того:

– сколько новых транспортных узлов должно быть построено в регионе;

– где эти новые транспортные узлы должны быть расположены;

– какие пропускные способности новые транспортные узлы и подъездные пути к ним должны иметь;

– какие схемы перемещения груза через новые и уже функционирующие транспортные узлы и подъездные пути к ним можно рассматривать как оптимальные для региона и для страны в целом;

– какие общие расходы связаны со строительством новых транспортных узлов и подъездных путей к ним и с обслуживанием всех элементов новой региональной инфраструктуры перевозок грузов;

– какой объем выручки планируется генерировать региональной инфраструктуре грузовых перевозок в виде налогов, разрешенных региональной администрацией быть предложенными полностью (или хотя бы частично) как ее финансовый вклад, например, в государственно-частное партнерство с потенциальными инвесторами был бы чрезвычайно полезным как федеральной, так и региональной администрации.

Такой инструмент позволил бы региональным администрациям анализировать эффективность региональных инфраструктур грузовых перевозок, как уже существующих, так и тех, которые будут разработаны для удовлетворения спроса на транспортные услуги в регионе. Поэтому, построение математической модели, лежащей в основе варианта инструментария для поддержки принятия решений, способного ответить на шесть перечисленных выше вопросов определения:

– оптимальных (с точки зрения региональной администрации) мест расположения новых грузовых транспортных узлов в регионе;

– общих расходов, связанных со строительством как новых транспортных узлов в выбранных (оптимальных) пунктах транспортной сети и подъездных путей к ним;

– ожидаемого дохода от функционирования разрабатываемых региональных инфраструктур грузового транспорта в любой период планирования представляет системную проблему для региональной администрации.

Решение этой проблемы для любого конкретного периода планирования предполагает наличие у региональной администрации информации об а) общем спросе на грузопотоки в каждом пункте происхождения груза, б) общей потребности в грузе в каждом пункте назначения груза, в) общем спросе на грузопотоки в регионе, г) стоимости затрат на строительство новых транспортных узлов, д) стоимости затрат на строительство новых подъездных путей к этим новым центрам, е) стоимости затрат на перевозку грузов между каждым местом происхождения груза и каждым пунктом назначения груза через каждый транспортный узел (как тот, который уже функционирует, так и тот, который планируется построить), и ж) стоимости обслуживания как транспортных узлов, так и подъездных путей к ним.

В зависимости от того, какие из значений перечисленных параметров считаются известными числами для лиц, принимающих решения, должны быть рассмотрены две ситуации.

Ситуация 1.

Конкретные числовые значения всех параметров, перечисленных в а) - ж), или только части этих параметров известны лицам, принимающим решения, на конкретный период планирования.

Ситуация 2

Только те области, к которым принадлежат значения всех параметров, перечисленные в а) – ж), известны лицам, принимающим решения, на конкретный период планирования.

Как показано в статье [194], если значения всех параметров, перечисленных в а) – ж), рассматриваются лицами, принимающими решения, как известные числа на конкретный период планирования (Ситуация 1, случай А), ин-

струментарий для поддержки принятия решений обеспечивает вышеупомянутое определение путем решения задачи математического программирования со смешанными переменными и с ограничениями и целевой функцией, имеющими линейную структуру. Эта задача может быть решена стандартными пакетами программ для решения задач оптимизации со смешанными переменными.

Если же значения всех параметров, перечисленных в а) - в), рассматриваются лицами, принимающими решения, как известные числа на конкретный период планирования (Ситуация 1, случай Б), то, как показано в статье [194], некоторая другая задача математического программирования со смешанными переменными, с линейными ограничениями и с целевой функцией, имеющими линейную структуру, должна быть решена.

Если для всех параметров, перечисленных в а) - ж), только области, к которым принадлежат их значения, известны лицам, принимающим решения (Ситуация 2), то как показано в статье [194], инструментарий поддержки принятия решений обеспечивает определение путем решения некоторой минимаксной задачи. Эта минимаксная задача имеет систему ограничений, обладающую линейной структурой, смешанные переменные и билинейную целевую функция двух векторных аргументов, одна из которых имеет только целочисленные координаты.

В статье [194] доказано, что эта минимаксная задача сводится к задаче математического программирования со смешанными переменными, с системой ограничений и с целевой функцией, имеющими линейную структуру, которые могут быть решены с помощью стандартных пакетов программ для решения задач математического программирования со смешанными переменными. Для обработки доступных данных для получения входной информации, необходимой как для расчетов в Ситуации 1 и в Ситуации 2, так и для графического изображения результатов расчетов может быть использовано другое стандартное программное обеспечение

В обоих случаях Ситуации 1, региональная администрация намерена потратить как можно меньше средств на (а) строительство обоих новых транспортных узлов и подъездных путей к ним, и (б) обеспечение технического обслуживания, как существующих транспортных узлов, так и подъездных путей к ним. Это те расходы, которые региональная администрация хотела бы видеть покрытыми частными инвесторами в рамках государственно-частного партнерства. В то же время региональная администрация предлагает свой финансовый вклад в это потенциальное партнерство. Этот вклад (или часть этого вклада) приходит в форме ожидаемого объема денежного потока, который будет создан налогами, которые будут уплачиваться в региональный бюджет грузовладельцами и грузоперевозчиками. Эти региональные налоги должны быть оплачены указанными клиентами новой грузовой транспортной инфраструктуры за использование ими транспортных узлов и подъездных путей к ним как элементов этой инфраструктуры.

Таким образом, общие расходы, которые сократились на ожидаемую сумму дохода от поступления вышеуказанных региональных налогов, должны быть минимизированы в обоих случаях Ситуации 1. Как показано в статье [194], соответствующие оптимизационные проблемы можно математически сформулировать в виде задач математического программирования со смешанными переменными и с системой ограничений, имеющей линейную структуру. В этих двух задачах математического программирования, как расходы, так и выручка математически описываются функциями смешанных переменных, имеющими линейную структуру. Решения этих задач определяют а) оптимальное расположение новых транспортных узлов, которые будут построены, наряду с их оптимальными пропускными способностями, б) оптимальный набор подъездных путей к этим новым транспортным узлам, которые будут построены, наряду с их типами и пропускными способностями и в) оптимальное распределение грузопотоков как по существующим транспортным узлам, так и по новым, соответствующие минимуму целевой функции в каждом случае, т.е. в Ситуации 1, Случай А и в Ситуация 1, Случай Б.

В Ситуации 2, целевой функцией в оптимизационной задаче по-прежнему является разность между расходами и ожидаемой суммой дохода, который будет получен в форме вышеуказанных региональных налогов. Тем не менее, в этой ситуации целевая функция оптимизационной задачи математически описывается функцией максимума функции алгебраической суммы билинейных и линейных функций векторных аргументов в конечномерных пространствах.

Компонентами векторных аргументов этих билинейных функций являются:

а) расходы, связанные со строительством новых транспортных узлов на каждом из мест их потенциального местоположения (значения стоимости строительства);

б) расходы, связанные со строительством подъездных путей к новым транспортным узлам в местах их потенциального местоположения (значения стоимости строительства);

г) расходы на техническое обслуживание, как новых транспортных узлов, так и подъездных путей к ним;

д) ожидаемые объемы грузопотоков, которые будут перемещаться между всеми транспортными узлами транспортной сети (как существующей, так и той, которая будет построена) и всеми местами отправления/назначения грузов;

е) булевы переменные, определяющие, должен ли новый транспортный узел определенной вместимости быть построенным в определенном пункте транспортной сети;

ж) булевы переменные, определяющие нужно ли строить новый подъездной путь определенного типа к новому транспортному узлу (или существующих подъездных путей достаточно для того, чтобы позволить

этому узлу функционировать на полную мощность) и пропускную способность каждого нового подъездного пути, который будет решено построить.

Непрерывные переменные из векторов, упомянутых в пунктах а) – г) этого а) – ж) списка, принадлежат выпуклым многогранникам, описываемым совместными системами линейных уравнений и неравенств. Естественно предполагать, что каждый из этих выпуклых многогранников, которому принадлежит каждая переменная вектора, является подмножеством параллелепипеда в соответствующем евклидовом пространстве. Булевы переменные, упомянутые в пункте ж) вышеупомянутого списка а) - ж), формируют два вектора, каждая из которых принадлежит единичному кубу (разным кубам для каждого из этих векторов) в соответствующем евклидовом пространстве.

Только финансирование работ, связанных с а) строительством новых транспортных узлов и подъездных путей к ним и б) управлением этими элементами новой региональной инфраструктуры грузовых перевозок составляет предмет настоящего отчета. Это, в частности, означает, что ни региональные пункты отправления / назначения груза, которые не используют существующие транспортные узлы и не будут использовать новые транспортные узлы, которые будут построены, ни подъездные пути к этим пунктам, которые в настоящее время используются, не будут рассматриваться. Конечно, эти пункты и эти подъездные пути являются элементами существующей грузовой транспортной инфраструктуры и останутся таковыми и в новой, которая будет разработана. Однако предполагается, что они не влияют на спрос на транспортные услуги в узлах региональной транспортной сети, которые будут соединены с транспортными узлами, входящими в грузовую транспортную инфраструктуру, подлежащую развитию. Таким образом, в рамках настоящего исследования развитие региональной транспортной инфраструктуры понимается как развитие только части этой инфраструктуры. Эта часть состоит только из пунктов зарождения грузов и назначения грузов, соединенных с существующими и с новыми транспортными узлами, как существующими дорогами, так

и новыми, которые будут построены для соединения с этими новыми транспортными узлами. Здесь некоторые из существующих подъездных путей, соединяют пункты, которые не являются частью региональной грузовой транспортной инфраструктуры, с существующими транспортными узлами могут, тем не менее, в настоящее время использоваться в качестве подъездных путей к существующим транспортным узлам (которые являются частью этой инфраструктуры) для перемещения грузов в них и из них. В рассматриваемой в [194] математической модели, возможности (пропускные способности) этих подъездных путей, которые считаются частью региональной грузовой транспортной инфраструктуры (которая будет развиваться) уменьшаются на объемы потоков грузов, перемещаемых вдоль этих дорог из пунктов транспортной сети, не являющихся транспортными узлами.

Еще одна особенность математической модели, представленной в [194], связана с учетом только «одного типа» грузов, следующих через все элементы региональной транспортной инфраструктуры, подлежащей разработке. Хотя, конечно, ожидается, что разные типы грузов должны быть перемещены через все инфраструктурные элементы, эта математическая модель работает с такими параметрами / переменными как емкости транспортных узлов, мощности подъездных путей к ним и потребности в грузовых услугах в пунктах отправления / назначения груза в регионе, в который будут перемещены грузы через эти транспортные узлы.

Однако эта модель не предназначена для расчета любых схем перемещения конкретных грузов через элементы региональной грузовой транспортной инфраструктуры, которая должна быть создана. Обычно на время принятия стратегических решений по финансированию создания и развития новых транспортных узлов и подъездных путей к ним, ни список таких грузов, ни объемы каждого груза из этого списка неизвестны точно. В то же время, а) ожидаемые суммарные объемы груза, который будет перемещен через каждый транспортный узел, создаваемый как часть региональной грузовой транспортной инфраструктуры, и б) ожидаемые общие объемы груза, который должен

быть перемещен через каждый подъездной путь к каждому из этих транспортных узлов, можно оценить. Здесь обе точные экспертные оценки этих общих объемов или, по крайней мере, областей, которым эти общие объемы могут принадлежать, удается получить. С помощью предложенной в [194] математической модели формулируются задачи оптимизации, решение которых позволяет региональной администрации получить оценку объема инвестиций, требуемых для создания новой региональной транспортной инфраструктуры грузовых перевозок в обоих приведенных выше случаях имеющейся информации об общих объемах грузов. Следует, однако, обратить внимание на то, что предложенная в [194] математическая модель может легко включить любую информацию о конкретных грузах, которую региональная администрация может в конечном итоге захотеть принять во внимание при расчете оценок объемов инвестиций, которые будут использованы в ходе переговоров с потенциальными инвесторами из частного сектора.

Математическая модель, предложенная в [194] и используемая при описании материала настоящего раздела, содержит ограничения (2) - (14) (в обозначениях статьи), связывающие непрерывные и дискретные (булевы) переменные. Именно эта модель является основой инструментария поддержки принятия решений, структура которого описана в [194] и который был опробован на нескольких наборах модельных данные, собранных авторами статьи с использованием открытого источника (см., например, https://github.com/ggfedin/Model_dataset). Описанию результатов тестирования именно этого инструментария и посвящена оставшаяся часть настоящего раздела отчета.

Целью тестирования было продемонстрировать, как инструментарий может быть использован в переговорах между региональной администрацией и потенциальными инвесторами. Именно, цель состояла в том, чтобы продемонстрировать, что для любого набора модельных данных, которые участники переговоров могут многократно менять, результаты расчетов, полученные с использованием этот инструментария, могут быть представлены в виде легко

читаемых таблиц и иллюстративных картинок, полезных для участников переговоров. Примеры таких таблиц представлены авторами в [194].

Важно подчеркнуть, что тестирование проводилось с целью демонстрации не более чем многообещающих возможностей, предлагаемых разработанным инструментарием для поддержки принятия решений, и оно не демонстрирует примеры его практического применения. Такой подход к представлению результатов тестирования связан с определенными трудностями в получении реальных данных, связанных с функционированием реальных региональных грузовых транспортных систем, по соображениям безопасности и конкуренции. В рамках естественного с точки зрения бизнеса и безопасности подхода, региональные администрации предпочитают иметь в своем распоряжении инструментарий, который позволил бы им самим проводить расчеты с любыми наборами данных, которые они могут решить использовать как реальные данные. Это особенно имеет место, если переговоры с потенциальными инвесторами из частного сектора по развитию, в частности, региональной инфраструктуры грузовых перевозок разворачиваются или планируются.

Тестирование, подробно описанное в [194], проводилось для модельного региона с 32 пунктами отправления / назначения грузов, «сконструированными» на основе информации, полученной из открытых источников. Предполагалось, что имеется два уже действующих транспортных узла (с номерами 1 и 2 в [194]) и восемь мест для потенциального размещения новых транспортных узлов (с номерами от 3 до 10 в [194]). Были рассмотрены два типа подъездных путей (железные дороги и автомобильные дороги) как к новым, так и к существующим транспортным узлам и предполагалось, что каждый тип подъездных путей может иметь две пропускные способности, из которых можно выбирать. В транспортные тарифы были включены три разные ставки налога ν (каждая из которых составила определенный процент от этих тарифов) и были рассмотрены три конкретных плановых периода (с продолжительностью ψ лет). Рисунок А1 из Приложения 3 статьи [194] показывает геогра-

фию расположения пунктов отправления / назначения груза в «сконструированный» регион, возможные места для новых транспортных узлов в этом регионе, а также расположение существующих транспортных узлов в регионе в некоторых условных географических координатах, которые используются в открытых источниках.

Были проведены расчеты по четырем вариантам задачи смешанного программирования («Ситуация 1, Случай А, Подслучай 1», «Ситуация 1, Случай А, Подслучай 2», «Ситуация 1, Случай В» и «Ситуация 1, Случай В, Подслучай 2») и по двум вариантам робастной (минимаксной) задачи оптимизации («Ситуация 2, Подслучай 1» и «Ситуация 2, Подслучай 2», см. Замечание 1 в разделе III статьи [194]). Для «Ситуации 1, Случай А, Подслучай 1» и для «Ситуации 1, Случай А, Подслучай 2», описанных в разделах I и III статьи [194], следующие данные получили конкретные значения: а) годовой потенциал транспортного узла для каждого из двух вариантов годовой мощности в каждом из восьми потенциальных мест расположения новых транспортных узлов, б) годовые мощности существующих двух транспортных узлов, в) годовая пропускная способность каждого нового подъездного пути к каждому новому транспортному узлу (которая соответствует годовой пропускной способности транспортного узла) для каждого из двух вариантов годовой пропускной способности транспортного узла, г) годовая пропускная способность каждой подъездного пути к каждому из двух существующих транспортных узлов, д) ожидаемый годовой спрос на грузовые услуги по каждому из вышеуказанных 32 пунктов отправления / назначения груза, е) ожидаемый суммарный годовой спрос на грузовые перевозки в регионе, ж) стоимость строительства нового транспортного узла для каждого из двух вариантов годовой пропускной способности узла, и) стоимость строительства автомагистрали ("хайвея") к новому транспортному узлу для обоих вариантов годовой пропускной способности автомагистрали, к) стоимость строительства железной дороги до нового транспортного узла для обоих вариантов годовой пропускной способности железной дороги, и л) расходы на техническое обслуживание автомагистрали для

обоих вариантов годовой пропускной способности автомагистрали и для железных дорог для обоих вариантов годовой пропускной способности железной дороги.

Были рассчитаны затраты на перевозку единицы объема груза между каждым из 32 пунктов отправления / назначения груза и между каждым из этих пунктов и каждым существующим и каждым новым транспортным узлом. Расчеты стоимости перевозки были проведены исходя из от: а) расстояния между вышеуказанными пунктами и узлами (рассчитанными с использованием Google Maps), б) средней стоимости транспортировки единичного объема груза на километр и с) скидки на стоимость перевозки, которая зависит от длины вышеупомянутого расстояния.

В Разделе III статьи [194] отмечается, что Случай 2 в Ситуации 1 и Случай 2 в Ситуации 2 отличаются от Случая 1 как в Ситуации 1, так и в Ситуации 2 только знаком перед третьим слагаемым в целевых функциях соответствующих оптимизационных задач, рассмотренных в разделе III из [194]. Именно, в Ситуации 1 для Подслучая 2 в Случае А и для Подслучая 2 в Случае В, целевая функция принимает вид (1') вместо (1) (см. [194]), в то время как в обеих соответствующих задачах оптимизации ограничения описываются соотношениями (2) - (14) (см. [194]).

Задачи (1) – (14) и (1') – (14) были решены с заданными и с рассчитанными параметрами, как было описано выше, и результаты расчета для «Ситуации 1, Случай А, Подслучай 1» и «Ситуации 1, Случай А, Подслучай 2» представлены в таблицах 1 и 2 из приложения 2 статьи [194], соответственно. Два примера расчета оптимальных размещений транспортных узлов и выбора их годовой мощности в результате решения задачи (1) – (14) для «Ситуации 1, Случай А, Подслучай 1» и в результате решения задачи (1′) – (14) для «Ситуации 1, Случай А, Подслучай 2» изображены на рисунке А2 и на рисунке А3 из Приложения 3 статьи [194], соответственно. Эти примеры были рассчитаны для конкретных значений двух параметров (ν=7; ψ=5), и результаты расчетов

представлены в строке 8 таблицы 1 и в строке 8 таблицы 2 в [194], соответственно.

Для робастной (минимаксной) задачи в Ситуации 2, описанной в разделе I статьи [194], с Подслучаями 1 и 2, описанными в разделе III статьи [194], приведены основанные на экспертных оценках, конкретные значения следующих параметров, которые были использованы для всех возможных новых транспортных узлов и для всех возможных новых подъездных путей к ним:

– максимальной и минимальной средней стоимости транспортировки единичного объема груза между каждым пунктом сети и новым (или существующим) транспортным узлом через каждый подъездной путь каждого типа для обоих вариантов годовой пропускной способности транспортного узла;

– максимальных и минимальных средних затрат на строительство нового транспортного узла в каждом из 8 мест для обоих вариантов ежегодной пропускной способности узла,

– максимальных и минимальных средних затрат на строительство нового подъездного пути каждого из двух типов (автомобильных дорог и железных дорог) до нового транспортного узла для обоих вариантов каждого типа подъездного пути (каждый вариант имеет годовую пропускную способность, соответствующую одному из двух вариантов годовой пропускной способности нового транспортного узла);

– максимальных и минимальных средних затрат на содержание каждого нового транспортного узла, для обоих вариантов годовой пропускной способности узла;

– максимальных и минимальных среднее расходов на содержание автомагистрали и железной дороги к каждому новому транспортному узлу для обоих вариантов годовой пропускной способности узла, для обоих вариантов каждого типа подъездного пути (каждый вариант имеет годовую пропускную способность, соответствующую одному из двух вариантов пропускной способности нового транспортного узла);

– максимальных и минимальных средних расходов на содержание автомагистрали и железной дороги к каждому существующему транспортному узлу.

Также были сделаны следующие допущения:

а) неизвестная стоимость перевозки единицы объема груза между пунктом сети и новым (или существующим) транспортным узлом через подъездной путь каждого конкретного типа не может превышать 130% и не может быть ниже 70% от текущей рыночной стоимости этой перевозки на расстояние между этим пунктом сети и этим транспортным узлом;

б) неизвестная стоимость строительства нового транспортного узла в любой точке сети для обоих вариантов годовой пропускной способности этого узла не может превышать 120% и не может быть ниже 91% от ее текущего значения (то есть того значения, которое соответствует существующей рыночной стоимости строительства);

в) неизвестная стоимость строительства автомагистрали до нового транспортного узла в любой точке для обоих вариантов годовой пропускной способности этого узла и обоих вариантов пропускной способности автомагистрали не может превышать 115% и не может быть ниже 85% от ее текущего значения (т.е. текущей рыночной стоимости);

г) неизвестная стоимость строительства железной дороги до новой транспортного узла в любой точке для обоих вариантов годовой пропускной способности этого узла и обоих вариантов годовой пропускной способности железной дороги не может быть выше 117,5% и не может быть ниже 82,5% ее текущего значения (т.е. текущей рыночной стоимости);

д) неизвестная стоимость обслуживания нового транспортного узла в любой точке для обоих вариантов годовой пропускной способности

этого узла не может превышать 125% и не может быть ниже 90% ее текущего значения (т.е. текущей рыночной стоимости);

е) неизвестные расходы на содержание автомагистралей и железной дороги до нового грузового транспортного узла в любой точке для обоих вариантов годовой пропускной способности этого узла, для обоих вариантов каждого типа подъездного пути (каждый вариант пути имеет годовую пропускную способность, соответствующую одному из двух вариантов годовой мощности транспортного узла), не могут превышать 150% и не могут быть ниже 50% их соответствующих (одному из двух вариантов ежегодной пропускной способности автомагистрали и ежегодной пропускной способности железной дороги) текущих значений (то есть, существующей в настоящее время рыночной стоимости);

ж) расходы на содержание существующего транспортного узла и расходы на техническое обслуживание подъездных путей к существующим транспортным узлам всех существующих типов являются известными числами, которые не изменяются в течение периода планирования.

Для «Ситуации 2, Подслучай 1», робастная (минимаксная) задача формулируется как задача (18) и для «Ситуации 2, Подслучай 2» как задача (18') в [194]. Задачи (18) и (18') были решены с использованием описанных выше входных данных. Результаты расчета для «Ситуации 2, Подслучай 1» представлены в таблице 5, и результаты расчета для «Ситуации 2, Подслучай 2» представлены в таблице 6 из приложения 2 статьи [194], соответственно.

Два примера расчета оптимальных размещений транспортных узлов, их пропускных способностей и оптимальных прикреплений пунктов отправления / назначения груза к транспортным узлам путем решения задачи (18) для «Ситуации 2, Подслучай 1» и путем решения задачи (18′) для «Ситуации 2, Подслучай 2» показаны на рисунке А6 и на рисунке А7 Приложения 3 статьи [194], соответственно. Эти примеры были рассчитаны для конкретных значений

двух параметров ($\nu = 7$; $\psi = 5$), и результаты расчетов представлены в строке 8 таблицы 5 и в строке 8 таблицы 6 статьи [194], соответственно.

В Ситуации 1 (в задаче (1) - (14) для Случая А и в задаче (1') - (14) для Случая Б и в Ситуации 2 (в задаче (18) для Подслучая 1 и в задаче (18') для Подслучая 2), системы ограничений и целевые функции в соответствующих задачах оптимизации были сформированы в соответствии с их описанием, представленным в Разделе III статьи [194]. Эти проблемы были решены с использованием пакета Intlinprog, являющегося частью интерактивной среды MatLab, установленной на личном ноутбуке. Ноутбук был оснащен 2,5-ГГц процессором Intel Core i5 и имел оперативную память 16-ГБ на платформе Windows. Оптимальные решения были получены менее чем за 0,5 секунды для задачи (1) - (14) и для задач (1') - (14) для девяти комбинаций значений ν и ψ в каждом из двух Подслучаев для каждой из этих двух проблем.

Оптимальные решения для задачи (18) и задачи (18') для девяти комбинаций значений ν и ψ в каждом из двух Подслучаев для каждой из этих двух проблем были получены менее чем за три секунды.

Для «Ситуации 1, Подслучай 1» и «Ситуации 2, Подслучай 1», при определенном наборе входных данных, решения задачи (1) – (14) и задачи (18) показывают, что годовой доход, который ожидается получить как результат функционирования региональных инфраструктур грузового транспорта, не покрывает необходимые расходы в течение одного года с момента начала функционирования всех новых объектов инфраструктуры. Тем не менее, за более чем годовой период функционирования обеих региональных инфраструктур грузового транспорта генерируется существенный доход, превышающий необходимые расходы. Вообще, на основе таких оценок (которые вполне ожидаемы для крупномасштабных проектов), региональная администрация может принять как минимум два стратегических решения (в модельной ситуации соответствует входным данным модели):

1) региональная администрация может предложить часть вышеупомянутого (существенного) дохода как ее финансовый вклад в государственно-частное партнерство в переговорах с потенциальными партнерами из частного сектора;

2) региональная администрация может принять решение не формировать никакого партнерства с частным сектором по проекту для обеспечения функционирования региональной инфраструктуры грузового транспорта, которую планируется развивать с какими-либо частными инвестициями или без них. Это может произойти, если а) ожидается, что проект будет приносить прибыль через относительно короткий период времени после всех того, как все объекты новой региональной инфраструктуры грузового транспорта начнут функционировать и б) региональная администрация может получить кредит, скажем, от банка на приемлемых условиях для покрытия, по крайней мере, расходов связанных с функционированием развитой инфраструктуры грузовых перевозок.

Согласно таблицам 1, 3 и 5 в [194], в модельной ситуации, соответствующей входным модельным данным, второе стратегическое решение может иметь место если:

а) оценки, полученные путем решения задачи (1) – (14) для «Ситуации 1, Случай А, Подслучай 1» и для «Ситуация 1, Случай В, Подслучай 1» на три года и на пять лет (при 7% и при 13% значения размера налога как в случае А, так и в случае В);

б) оценки, полученные при решении задачи (18) для «Ситуации 2, Подслучай 1» на три года и на пять лет (при 13% от значения размера налога).

В дополнение к этому, в зависимости от условий потенциального кредита, могут быть определенные комбинации обеих стратегий. Далее определение того, сколько брать в долг и сколько просить частных инвесторов внести

как их вклад может потребовать использования дополнительных математических моделей и методов. Наконец, другие стратегии могут быть сформированы на основе приведенных выше оценках ожидаемых финансовых результатов функционирования проекта.

В то же время региональная администрация должна иметь в виду, что конкретные результаты расчетов всегда получаются для определенного набора данных, и изменение данных может привести к расчету стратегии, которая может показаться более перспективной. Кроме того, конкретные результаты расчетов могут указывать на то, что значения некоторых параметров, отражающих условия неопределенности должны быть пересмотрены. Например, границы, в пределах которых могут варьироваться ожидаемые объемы грузопотоков, могут быть такими параметрами. Таким образом, оценки, которые могут быть рассчитаны с использование предлагаемого инструмента поддержки принятия решений может обеспечить определенную гибкость региональной администрации при выборе ею своей финансовой стратегии развития региональной инфраструктуры грузового транспорта.

Раздел III в [194] описывает как частные инвесторы, участвуя в переговорах с региональной администрацией о потенциальных инвестициях в а) развитие новой региональной инфраструктуры грузовых перевозок и б) обеспечение транспортных услуг в рамках новой региональной инфраструктуры грузовых перевозок могут определить, возможно ли извлечение ими выгоды из соответствующих инвестиций с использованием предложенного инструмента поддержки принятия решений. Они могут сделать это, решая задачу (1') - (14) и задачу (18'). Однако для решения задачи (18′) им необходимо знать функцию, описывающую их расходы, связанные с предоставлением транспортных услуг в рамках новой транспортной инфраструктуры.

Ясно, что в a) «Ситуации 1, Случай A, Подслучай 2», b) «Ситуации 1, Случай B, Подслучай 2» и c) «Ситуации 2, Случай 2 », региональная администрация может повлиять на общие расходы ее потенциальных партнеров из частного сектора путем изменения размера налогов. Результаты расчетов при

конкретном наборе входных данных, представленные в приложении 2 в [194], показывают, что доля налогов в общем объеме расходов оказывается относительно небольшой для однолетних проектов. Самая низкая доля для этих проектов составляет около 1% (как показывает результат решения задачи (1′) - (14) для «Ситуации 1, Случай B, Подслучай 2» для значений параметров $\psi = 1$, $\nu = 1$), а самая высокая доля, около 12% (как показывает результат решения задачи (1′) - (14) для «Ситуации 1, Случай A, Подслучай 2» для значений параметров $\psi = 1$, $\nu = 13$). В то же время, для пятилетних проектов, самая низкая доля составляет около 3% (как показывает результат решения задачи (18′) для «Ситуации 2, Подслучай 2» для значений параметров $\psi = 5$, $\nu = 1$), а самая высокая доля составляет около 31% (как показывает результат решения задачи (1′) - (14) для «Ситуации 1, Случай A, Подслучай 2» для значений параметров $\psi = 5$, $\nu = 13$).

На основании этих результатов расчета (которые, как следует еще раз подчеркнуть, были проведены для конкретного набора только лишь модельных данных путем решения задачи (1′) – (14) и задачи (18′)), можно сделать вывод, что долгосрочные проекты по развитию региональной инфраструктуры грузовых перевозок могут при определенных обстоятельствах быть более чувствительными к изменению в налоговой стоимости, чем краткосрочные.
Подробное обсуждение полученных результатов и мнение авторов о роли разработанного ими инструментария по поддержке принятия решений, в частности, в транспортных и экономических системах, содержится в [194].

3.2.3 Выявление системных проблем размещения объектов и привлечения частных инвестиций

Постоянный рост населения мира делает крупномасштабные энергетические проблемы одними из самых важных проблем, с которыми сталкивается человечество. Создание и использование систем государственного и муниципального управления в области энергетики свидетельствует об осознании правительствами разных стран своей ответственности за решение этих проблем,

что подразумевает соответствующее их финансирование, которое, на сегодняшний день, остается недостаточным, лишая миллионы людей на планете доступных энергетических услуг [197]. Поиск подходящего финансирования для любого конкретного крупномасштабного энергетического проекта по-прежнему представляет значительные трудности даже тогда, когда необходимость реализации этого проекта выглядит очевидным для тех, кто принимает соответствующие организационные и финансовые решения. Относительно новые формы финансового сотрудничества между государственным и частным секторами экономики часто составляют единственную основу для финансирования этих проектов [198]. Концессии, долгосрочные договоры лизинга территории, и государственно-частные партнерства являются такими формами сотрудничества, превращающими региональную и местную общественность в основных игроков в инициации, реализации и управлении этими проектами, и делают крупномасштабные энергетические проекты мало зависящими от федеральных бюджетов.

Существует по крайней мере шесть основных системных проблем, связанных с разработкой и реализацией крупномасштабные проектов по привлечению частных инвестиций для разработки новых и модернизации существующих энергетических систем, исследование которых представляет существенный интерес для региональных государственных администраций.

Источники энергии в региональных энергетических системах, включая новые источники, являются одной из таких системных проблем. Согласно экспертным оценкам, наличие мест для хранения традиционных источников ископаемого топлива ограничено и эти лимиты скоро будут превышены [199], что способствует интересу к новым источникам энергии. Уже сегодня ядерная энергетика, и энергия ветра, солнца и морского прибоя обеспечивают почти 20% всей энергии, потребляемой в мире [200]. Твердые бытовые отходы в крупных городах и тающий лед дают еще два примера многообещающих новых источников энергии, которые интересуют государственные администра-

ции. Включение перспективных источников энергии в настоящее время функционирующие электрические сети в конкретных географических регионах имеет значение и с точки зрения проблем охраны окружающей среды. Поэтому государственные администрации ожидают от исследователей в области энергетики количественных экономических оценок перспективных областей энергетики и направлений развития существующих электрических сетей с учетом новых источников энергии [201]. В частности, в области анализа стоимости разработки, хранения и транспортировки энергии, которые позволили бы оценить, как связанные с ними затраты, так и возможную прибыль от эксплуатации соответствующих энергетических систем, создавая тем самым основу для переговоров с частным сектором об условиях совместного финансирования создания этих систем.

Технологии получения энергии из ее источников представляют интерес для всех государственных администраций, так как эти технологии влияют на окружающую среду, а охрана окружающей среды входит в их обязанности [202]. Производительность отдельных технологий, которые используются для получения энергии из ее источников в регионе также имеет значение [203]. Количественные закономерности, описывающие как уже развернутые технологии получения энергии, так и те, использование которых ожидается в ближайшем будущем, могут повлиять на окружающую среду, важны государственного управления [204]. Кроме того, эти количественные закономерности, позволяют государственным администрациям и общественности количественно оценить, как эти технологии влияют на стоимость соответствующих услуг для населения. В частности, эти закономерности помают им: а) найти оптимальное распределение финансовых ресурсов, связанных с реализацией крупномасштабных энергетических проектов, и б) оценить объем инвестиций, которые необходимо получить от частного сектора с этой целью.

Распределение и транспортировка энергии. Источники энергии для потребителей интересуют все государственные администрации, поскольку либо территории, на которых эти источники энергии находятся, либо территории,

через которые конкретные виды энергии могут быть доставлены своим клиентам (либо и те, и другие территории), находятся под юрисдикцией соответствующих региональных органов власти [205]. Причем, как правило, все оптовые энергетические рынки имеют отдельные системы транспортировки энергии, что превращает каждую такую систему в естественную монополию, регулируемую правительственными учреждениями [206]. Это объясняет то, почему: а) правильное прикрепление всех доступных источников энергии к ее потребителям с указанием того, как много энергии должно быть доставлено из каждого источника энергии к каждому конкретному потребителю и б) отыскание наилучшего способа доставки этой энергии к потребителям так интересует государственные администрации [207]. Все такие прикрепления определяют как структуру энергетических сетей, так и расходы, связанные с транспортировкой единицы объема энергии каждого типа с помощью каждого возможного способа доставки. В свою очередь эта информация может помочь государственным администрациям договориться о формировании государственно-частных партнерств, с частным бизнесом в энергетическом секторе вообще и в области распределения и транспортировки энергии в частности. Таким образом, системы для расчета оптимальных заданий и оптимальных маршрутов доставки энергии клиентам в определенный регион представляют большой интерес для государственных администраций, управляющих конкретными географическими районами.

Системы энергоснабжения включают в себя энергетические сети, которые должны быть количественно проанализированы как на этапе их создания, так и на этапе их развития и модернизация [208]. Несколько условий, при которых анализируются функции энергетических сетей, включают в себя учет: а) пиковых требований и пиковых нагрузок; б) неожиданных неэкономических обстоятельств, в том числе погодных условий; в) неожиданных экономических условий, окружающих функционирование энергетических сетей; г) разных пропускных способностей границ сетей; и д) нестабильности энерге-

тического рынка. Обеспечение надежного энергоснабжения как населения региона, так и находящихся там государственных и частных предприятий через существующие и планируемые энергетические сети, обычно рассматривается жителями региона как один из критериев, отражающих качество административного управления [209]. В то же время многое зависит от того, как энергетическая сеть разработана, и как она развивается со временем. Здесь математические модели с учетом вышеуказанных условий и неопределенности (при которой функционирует большинство энергетических сетей) играют принципиальную роль в создании эффективных систем поддержки принятия решений в области энергетики, которые помогают государственному администрацию анализировать финансовые и социальные аспекты каждого плана модернизации регионов, находящихся под их юрисдикцией [210].

Системы хранения энергии представляют собой важный инструмент государственного управления для обеспечения бесперебойного энергоснабжения всех потребителей энергии, проживающих и / или работающих в регионе. Помимо территорий, где эти системы хранения могут быть расположены, эти системы требуют надежных и точных управленческих решений. Эти решения подразумевают: а) вычисление объемов энергии, которые эти системы могут хранить [211], б) определение структуры включения этих систем в уже функционирующие энергетические сети [212] и в) выбор стратегии использования систем хранения энергии [213]. Хотя все указанные три компоненты управленческих решений являются прерогативами локальных властей, некоторые общие стратегические решения, связанные с управлением системами хранения энергии должны быть приняты в каждой стране. Для эффективного управления системами хранения энергии, как на местном, так и на федеральном или на национальном уровне требуются инструменты для количественного анализа возникающих здесь проблем. Эти инструменты должны позволять: а) получать краткосрочные и долгосрочные оценки оптимальных размеров систем хранения и б) определить стратегии их пополнения энергией. Хотя математические модели, лежащие в основе таких инструментов, различаются, все они

должны учитывать рассмотрение, как масштабы этих систем, так и различные неопределенности при которых они функционируют. Модели математического программирования с линейными ограничениями выглядят наиболее перспективными с этой точки зрения, так как проблемы управления хранением энергии, сформулированные на основе любых моделей, должны эффективно решаться при различных неопределенностях. В настоящее время это может быть сделано с помощью мощных стандартных программных пакетов для решения задач математического программирования с линейными ограничениями. Также следует учитывать, что управление системами хранения электроэнергии должны осуществляться вместе с управлением доступными источниками энергоснабжения определенного региона или определенных регионов [214]. Таким образом, какими бы ни были математическими модели, лежащие в основе конкретных систем управления хранением энергии, эти модели должны учитывать наличие: а) источников энергии, б) собственно систем хранения энергии и с) финансовых и людских ресурсов, необходимых для реализации всех управленческих решений.

Эффективное использование энергии в населенных пунктах для общественных нужд является прямой обязанностью каждой системы государственного управления, которое, однако, может быть осуществлено только в рамках доступных источников энергии и имеющихся или предполагаемых к построению систем хранения энергии [215]. Инструменты, позволяющие оценивать ожидаемый общественный спрос на все виды энергии исходя из существующей демографических изменений в соответствующих географических областях представляют значительный интерес, как федерального правительства, так и для соответствующих региональных администраций. Тот же интерес для любой системы государственного управления представляют инструменты, способные прогнозировать существующие и ожидаемые тенденции в использовании определенных видов энергии в конкретных географических областях [216]. Например, обеспечение электричеством площадей, включая парки, скверов, дорог, мостов, автобусов, железнодорожных станций, аэропортов и

т.д.; и б) всех домов в жилых районах является примером проблемы, требующей решение в каждой географической области на основе вышеупомянутых прогнозов демографических изменений. Обеспечение водителей автозаправочных станций, где они могут заправлять бензином свои автомобили, как в городах, так и вдоль автомагистралей, дает еще один пример общественных требований, которые должны быть выполнены [217]. Хотя соответствующие заправочные станции обычно управляются частными компаниями, региональные администрации несут ответственность за стратегии развития цепочек этих станций. Для реализации этих стратегий, государственные администрации выдают соответствующие разрешения и / или лицензии для этих частных компаний. Количественный расчет таких стратегий, особенно в соответствии с существующими тенденциями проникновения электрических автомобилей в повседневную жизнь городов и пригородов, подразумевающий проведение тщательного анализа рассматриваемой проблемы, может быть сделан только с использованием соответствующих математических моделей.

Даже этот краткий взгляд на рассмотренные энергетические проблемы показывает, что государственные администраторы: а) должны иметь дело с этими проблемами, и б) могут иметь свое видение того, что им нужно для того, чтобы успешно решать эти проблемы. Таким образом, ожидания государственных администраций могут рассматриваться как определяющие рамки, в которых энергетические проблемы, требующие для своего решения реализации крупномасштабных проектов в отрасли, должны быть рассмотрены. Здесь сама природа государственных услуг обеспечивает уникальные возможности для приложений принятия решений на практике. В то же время, этой области присущи нестандартные вызовы, связанные с необходимостью предложения решения энергетических проблем для широкого спектра населения всего мира. Именно, эти уникальные возможности определяются в основном экономическими причинами, так как все государственные администрации имеют дело с энергетическими проблемами в рамках ограниченных бюджетов. Широкий спектр свободы в предложении партнерских отношений частному сектору в

различных условиях, которые могут выбирать государственные администрации в рамках закона, также характеризует такие уникальные возможности. В то же время, крупные энергетические проекты, как правило, финансируются деньгами налогоплательщиков и поскольку реализация этих проектов непосредственно влияет на жизнь людей, на государственных администрациях лежит огромная ответственность за реализацию каждого такого проекта, чтобы обеспечить широкий круг общественных интересов с ответственностью и профессионализмом.

С точки зрения автора отчета, применение количественных методов принятия решений является наиболее перспективным ответом на указанные выше вызовы системного характера, в том числе связанные с формированием государственно-частного партнерства для выполнения крупномасштабных проектов.

Некоторые проблемы из перечисленных шести классов системных проблем и примеры применения количественных методов принятия решений рассмотрены в том же самом специальном выпуске журнала, в котором опубликованы статьи [194, 195, 196].

3.2.4 О формализации и математическом обосновании утверждений

В отчете 2017 г. [218] рассматривалась содержательная постановка задачи об отыскании оптимального размещения цепочки зарядных станций для электромобилей. Для этой задачи была предложена математическая модель, позволяющая сформулировать эту задачу в минимаксной постановке при некоторых предположениях приведенных в этом отчете. В 2018-2019 гг. эта модель была модифицирована, в частности, за счет ослабления ряда указанных предположений, и эта модель, имеющая специальную структуру, приведена в [196].

Эта структура позволяет формализовать рассматриваемую задачу как задачу математического программирования, которая может быть эффективно

решена с помощью стандартных пакетов программ. Структура модели позволяет оценить объем инвестиций, необходимых для развития сети в условиях неопределенности на: а) экспертные оценки спроса на электроэнергию от электромобилей на каждой зарядной станции цепи и б) цены на электричество в электрической сети (сетях), поставляющей (поставляющих) электричество в цепь зарядных станций.

Эффективность решения формализованной проблемы, в частности, подразумевает возможность получения количественных оценок размера инвестиций даже когда: а) эта проблема является крупномасштабной (т.е. имеет большое число переменных и ограничений) и б) некоторые из переменных задачи могут быть целыми числами.

Кроме того, решение формализованной задачи, которая формулируется на основе предложенной модели, может быть найдено за короткий период времени, и его отыскание не требует разработки каких-либо новых математических методов и алгоритмов. Именно, это решение может быть найдено с использованием стандартных пакетов программного обеспечения, позволяющих быстро и точно решать крупномасштабные задачи. Два известных пакета программ CPLEX (https://www.ibm.com/analytics/cplexoptimizer) и MILP (https://www.mathworks.com/help/optim/ug/ intlinprog.html) являются примерами таких пакетов программного обеспечения для решения задач математического программирования со смешанными (непрерывными и дискретными) переменными, включая задачи линейного, целочисленного и смешанного программирования.

В предлагаемой математической модели специальной структуры присутствуют нелинейные функции, что объясняется вышеупомянутой неопределенностью экспертных оценок спроса на электроэнергию и цен на электроэнергию. Именно, эти оценки сделаны тогда, когда ни стоимость обслуживания конкретного электрического транспортного средства определенного типа при зарядке на определенной станции в определенный период времени, ни число электромобилей каждого типа, которые будут обслуживаться на этой

станции в этот период, заранее точно неизвестны. В результате математическая модель, которая будет использоваться для формализации рассматриваемой задачи, оказывается нелинейной.

При разработке математической модели для этой задачи делается предположение относительно структуры спроса на электроэнергию. Именно, предполагается, что эта структура та же самая на каждой зарядной станции цепи в пределах каждых 24 часов – с полуночи до следующей полночи - в течение каждого из четырех сезонов (зима, весна, лето и осень) хотя, как правило, эти структуры разные для разных сезонов. Однако, предлагаемый подход к моделированию задачи не ограничивается четырьмя «сезонами» в пределах каждого из которых структуры спроса на электроэнергию одинаковы. Число таких «сезонов» может быть любым, от четырех до 365/366 каждый год.

Следует заметить, что в [196] рассматривается только случай, в котором предполагается, что все электричество поступает на все зарядные станции цепи из электрической сети (или из разных электрических сетей на разные станции) напрямую. Однако, как упоминается во Введении в [196], подход к расширению предложенной математической модели, с целью охватывать, в частности, случаи, когда электричество к зарядным станциям может также поступать от а) электрической сети (или других электрических сетей) через системы хранения электроэнергии, и б) возобновляемых источников энергии (опять же, как напрямую, так и через системы хранения электроэнергии).

В [196] приводится описание переменных и ограничений математической задачи, являющейся формализованным описанием задачи, исследовавшейся в 2018-2019 гг. В ней показано, что эта задача может быть сформулирована как задача отыскания максимина функции, являющейся суммой разности двух билинейных функций, алгебраической суммы трех линейных функций и скалярного произведения фиксированного вектора и вектора, все компоненты которого являются неотрицательными целыми числами, на прямом произведении нескольких выпуклых многогранников и множества, являющегося под-

множеством выпуклого многогранника, образованным точками, все координаты которых являются неотрицательными целыми числами. Основной математический результат, полученный в работе над рассматриваемой задачей и опубликованный в [196], состоит в доказательстве возможности сведения указанной задачи отыскания максимина к задаче максимизации функции, имеющей линейную структуру, при ограничениях также имеющих линейную структуру, связывающую переменные конечномерных пространств, причем часть этих переменных принимает только неотрицательные целые значения. Полученная в результате этого сведения задача математического программирования со смешанными переменными может быть решена с помощью пакетов CPLEX и MILP. Формулировка и доказательство соответствующей теоремы приведены в [196].

4 Исследования моделей анализа политических систем

В этом разделе отчета приведены результаты исследований моделей в политической сфере. Эти исследования проходили в следующих направлениях:

– анализ значимости позиций партий для голосования по ним методами теории монотонных мер;

– исследования поляризованности и влияния в политических системах.

4.1 Анализ значимости позиций партий для голосования по ним методами теории монотонных мер

4.1.1 Введение и постановка задачи исследования

Распространённой системой выборов депутатов в парламент является выбор по партийным спискам. Для информирования граждан, в некоторых странах проводятся интернет-опросы, в которых политические партии перед выборами в парламент озвучивают свою точку зрения на некоторые актуальные для избирателей вопросы. Например, такой сервис хорошо развит в Германии [219]. Данные подобных сервисов и их сопоставление с итогами последующих выборов анализируются как политологами, так и специалистами-математиками [220]. Например, представляют интерес следующие вопросы: 1) значимы ли позиции партий по вопросам, сформулированных в интернет-опросах (а они формируются по результатам голосования за них избирателей) для результатов последующего голосования в парламент? 2) как выявить группы действительно значимых вопросов? 3) как оценить политическую неоднородность общества? 4) как оценить согласованность или конфликтность позиций партий по значимым вопросам? Последний вопрос важен, например, для оценки возможности создания коалиций.

Все эти вопросы широко обсуждаются в политологической литературе. Например, в [221, 222] анализируется согласованность голосований между

фракциями или внутри фракций. Существуют различные (прежде всего теоретико-игровые) модели формирования и функционирования фракций и коалиций [223, 224, 225, 226, 227]. Возможность создания коалиций часто рассматривается при анализе коалиционной устойчивости [228] или структурной устойчивости [229].

Мы приводим результаты другого подхода. А именно, с точки зрения анализа данных приведенные выше вопросы имеют функционально-множественный характер – значимость необходимо рассматривать не для отдельных вопросов, а на подмножествах вопросов, поскольку возможен «синергетический» эффект значимости некоторых групп вопросов. Поэтому для ответа на этот вопрос необходимо использовать соответствующий инструментарий. Таким инструментарием, на наш взгляд, может быть теория монотонных мер [230], которая учитывает «синергетический» эффект отдельных подмножеств вопросов, а также некоторые методы теории функций доверия [231, 232].

В этом подразделе отчета приведены результаты такого исследования на примере интернет-опросов в Германии перед выборами в Бундестаг в 2013г. [233]. В исследовании рассмотрены и решены задача нахождения наиболее значимых для голосования групп вопросов, задача оценки влияния позиций партий на результат голосования, задача оценки конфликтности позиций партий по значимым вопросам, задача оценивания политической неоднородности общества.

4.1.2 Некоторые сведения из теории неаддитивных мер, интегралов и функций доверия

Пусть $X = \{x_1, ..., x_n\}$ – некоторое конечное множество (например, множество альтернатив, критериев, признаков и т.д.), 2^X – множество всех подмножеств из X. Тогда нормализованной монотонной мерой [230] (неаддитивной мерой, нечеткой мерой, емкостью) на X называется функция множеств $\mu : 2^X \to [0,1]$, которая удовлетворяет условиям: 1) $\mu(\varnothing) = 0$, $\mu(X) = 1$;

2) $\mu(A) \le \mu(B)$, если $A \subseteq B$, $A, B \in 2^X$. В задачах принятия решений или выделения признаков значение $\mu(A)$ характеризует важность множества признаков (критериев, альтернатив и т.д.) A для решаемой задачи.

Дискретным интегралом Шоке от неотрицательной функции $f : X \to \mathbb{R}_+$ по монотонной мере μ на X называют величину [234]

$$C_\mu(f) = \sum_{i=1}^{n} \Big(f(x_{\sigma(i)}) - f(x_{\sigma(i-1)}) \Big) \mu\Big(A_{\sigma(i)}\Big),$$

где $\sigma(i)$ – такое упорядочение индексов $\{1,\ldots,n\}$, что $0 = f(x_{\sigma(0)}) \le f(x_{\sigma(1)}) \le \ldots \le f(x_{\sigma(n)})$, $A_{\sigma(i)} = \{x_{\sigma(i)},\ldots,x_{\sigma(n)}\}$, $i = 1,\ldots,n$. Интеграл Шоке, как агрегирующий оператор, находит многочисленные применения в задачах распознавания образов [235, 236] и information fusion: многокритериальном принятии решений [230, 237], выделении особенностей (feature extraction) [238], слияния данных в мультисенсорных системах [239, 240] и т.д.

В ряде задач дискретный интеграл Шоке удобно представлять в терминах преобразования Мёбиуса от монотонной меры μ. Преобразованием Мёбиуса функции множеств μ на X называют функцию множеств $m^\mu : 2^X \to \mathbb{R}$, которая определяется как

$$m^\mu(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} \mu(B), \; A \in 2^X .$$

С другой стороны, по преобразованию Мёбиуса можно восстановить исходную функцию множеств

$$\mu(A) = \sum_{B \subseteq A} m^\mu(B), \; A \in 2^X .$$

Значение $m^\mu(A)$ характеризует важность множества A безотносительно его связи с другими его подмножествами.

Монотонная мера $\mu : 2^X \to [0,1]$ называется k-аддитивной, $k \in \{1,...,n\}$ [241], если $m^\mu(A) = 0$ для всех $A \in 2^X$, $|A| > k$ и существует такое $A \in 2^X$, $|A| = k$, что $m^\mu(A) \neq 0$.

В терминах преобразования Мёбиуса интеграл Шоке примет вид [242]

$$C_\mu(f) = \sum_{A \subseteq X} m^\mu(A) \min_{x \in A} f(x) \cdot$$

Если $m^\mu(A) \geq 0$ для всех $A \in 2^X$, то монотонная мера μ называется мерой (функцией) доверия. Двойственная к функции доверия мера $\bar{\mu}(A) = 1 - \mu(\neg A)$, $A \in 2^X$ называется мерой (функцией) правдоподобия. Функции доверия и правдоподобия являются одним из основных объектов исследования теории функций доверия (теории свидетельств, теории Демпстера-Шейфера) [231, 232].

В теории функций доверия предполагается, что изначально задана или статистически оценена функция масс $m : 2^X \to [0,1]$, которая определяет первоначальную информацию о свидетельстве. Затем по функции масс строятся функции доверия Bel и правдоподобия Pl:

$$Bel(A) = \sum_{B \subseteq A} m(B), \; Pl(A) = \sum_{B \cap A \neq \varnothing} m(B), \; A \in 2^X .$$

Вклад отдельных альтернатив в общую важность подмножеств альтернатив можно оценить с помощью значений Шепли [243], которые широко используются в теории кооперативных игр и вычисляются по формуле

$$v_i = \sum_{A \subseteq X,\, x_i \in A} \frac{(n - |A|)!(|A| - 1)!}{n!} \left(Bel(A) - Bel(A \setminus \{x_i\}) \right) \cdot$$

Вектор $\mathbf{v} = (v_i)_{i=1}^n$ называется вектором Шепли, $\sum_{i=1}^n v_i = Bel(X) = 1$. Известно [244], что для функций доверия вектор Шепли совпадает с так называемым пигнистическим преобразованием (pignistic transformation):

$$v_i = \sum_{A \subseteq X,\, x_i \in A} \frac{m(A)}{|A|}.$$

В теории функций доверия множество $A \in 2^X$ называется фокальным элементом свидетельства, если $m(A) > 0$. Пусть A – множество всех фокальных элементов свидетельства. Тогда пару $F = (A, m)$ называют телом свидетельств. Одним из важных вопросов в теории свидетельств является оценивание конфликта между несколькими свидетельствами (внешний конфликт), заданными на множестве X, а также оценивание внутреннего конфликта каждого свидетельства. Исторически первой и наиболее популярной мерой внешнего конфликта является конъюнктивная мера конфликта, которая для двух тел свидетельств $F_1 = (A_1, m_1)$ и $F_2 = (A_2, m_2)$ на X вычисляется по формуле

$$K = K(F_1, F_2) = \sum_{A \cap B = \varnothing} m_1(A) m_2(B). \tag{4.1}$$

Нетрудно видеть, что $K \in [0,1]$ и значение $K = 1$ соответствует абсолютной конфликтности двух тел свидетельств: $A \cap B = \varnothing$ для любых $A \in A_1$ и $B \in A_2$. Эту меру мы будем использовать ниже для оценивания конфликтности позиций партий. Исследования аксиоматики, свойств и способов задания мер внешнего конфликта можно найти в [245, 246, 247, 248].

Внутренний конфликт свидетельства характеризует степень его неоднородности. В частности, если свидетельство содержит непересекающиеся фокальные элементы с большими массами, то это говорит о большом внутреннем конфликте такого свидетельства. Существуют различные способы оценивания внутреннего конфликта [249, 250, 251]. Например, мерой внутреннего конфликта тела свидетельств $F = (A, m)$ можно считать величину $K(F, F)$ (так называемый автоконфликт), где мера K вычисляется по формуле (1). С другой стороны, меру внутреннего конфликта можно считать небольшой, если

$\Omega_A = \bigcap_{A \in A} A \neq \varnothing$. В этом случае $Pl(x) = \sum_{A \in A\ : x \in A} m(A) = 1 \quad \forall x \in \Omega_A$. Основываясь на этом наблюдении в [249] была введена мера внутреннего конфликта, как $K_{\text{int}}(F) = 1 - \max_x Pl(x) = \min_x Bel(X \setminus \{x\})$. Эту величину в данной работе мы будем использовать для оценивания внутреннего конфликта.

4.1.3 Постановка основных задач исследования

Пусть имеется r партий, которые на выборах набрали $p_1, ..., p_r$ голосов соответственно. Кроме того, имеется список из n вопросов $N_n = \{1, ..., n\}$, по которым все партии до выборов озвучили свои позиции, 2^{N_n} – множество всех подмножеств списка вопросов. На каждый из вопросов партия может ответить либо «да», либо – «нет». Поэтому для любого множества вопросов $A \in 2^{N_n}$ существует $2^{|A|}$ подмножеств с учетом различных ответов на вопросы из A. Пусть 2^A – множество всех таких ответов на вопросы из A, а $2^{\overline{N_n}} = \bigcup_{A \in 2^{N_n}} 2^A$ – множество всех подмножеств списка вопросов с учетом различных ответов, $\left|2^{\overline{N_n}}\right| = 3^n$. Каждому множеству $A \in 2^{N_n}$ можно поставить в соответствие вектор $a = (a_1, \ldots, a_n)$, $a_i = 1$: если $x_i \in A$ и $a_i = 0$ в противном случае. Каждому множеству $A \in 2^{\overline{N_n}}$ можно поставить в соответствие вектор $a = (a_1, \ldots, a_n; a_{n+1}, \ldots, a_{2n})$: если $x_i \in A$ и ответ был «да», то $a_i = 1$; если $x_i \in A$ и ответ был «нет», то $a_{n+i} = 1$ и $a_i = 0$ в противном случае.

4.1.3.1 Нахождения наиболее значимых для голосования групп вопросов и оценка политической неоднородности общества

Необходимо определить значимость того или иного подмножества вопросов для голосования, т.е. необходимо найти функцию масс $m : 2^{N_n} \to [0,1]$,

где значения $m(A)$ характеризуют доли голосов избирателей, считающих вопросы множества A значимыми, $\sum_A m(A)=1$. Вместо функции множеств m удобней искать функцию $\bar{m}:2^{\overline{N_n}}\to[0,1]$. Тогда $\sum_{B\in 2^A}\bar{m}(B)=m(A)$, $A\in 2^{N_n}$.

Пусть

$$v_{k,i}=\begin{cases}1, \text{ если } i\in\{1,...,n\} \text{ и } k\text{-я партия ответила "да" на } i\text{-й вопрос},\\ 1, \text{ если } i\in\{n+1,...,2n\} \text{ и } k\text{-я партия ответила "нет" на } (i-n)\text{-й вопрос},\\ 0, \text{ в противном случае},\end{cases}$$

$$k=1,...,r\ ,\ i=1,...,2n$$

и $V_k=(v_{k,1},...,v_{k,n};v_{k,n+1},...,v_{k,2n})$ – профиль ответов на вопросы k-й партии. Нетрудно видеть, что для профиля k-й партии верно равенство $v_{k,s}+v_{k,n+s}=1$ для всех $s=1,\ldots,n$. Например, для $n=4$ вопросов профиль $V=(1,0,0,1;0,1,1,0)$ означает, что партия ответила «да» на 1-й и 4-й вопросы и «нет» на 2-й и 3-й вопросы.

Тогда величина

$$C(V_k)=\sum_{A\in 2^{\overline{N_n}}}\bar{m}(A)\min_{i\in A}\{v_{k,i}\} \tag{4.2}$$

характеризует потенциальное число избирателей k-й партии. Необходимо найти такие значения $\{m(A)\}$, чтобы $\dfrac{C(V_k)}{C(V_j)}\approx\dfrac{p_k}{p_j}$ для любых пар индексов $k,j\in\{1,...,r\}$. Для последнего условия можно рассмотреть некоторые критерии минимизации. Например, $R=\sum_i\left(p_{i+1}C(V_i)-p_iC(V_{i+1})\right)^2$, считая, что $p_{r+1}=p_1$, $V_{r+1}=V_1$.

Таким образом, необходимо найти значения $\left\{\bar{m}(A):A\in 2^{\overline{N_n}}\right\}$, удовлетворяющие условиям

$$\sum_{A\in 2^{\overline{N_n}}} \bar{m}(A)=1,\ \bar{m}(A)\geq 0 \text{ для всех } A\in 2^{\overline{N_n}} \tag{4.3}$$

и минимизирующие критерий R. Это задача квадратичного программирования. Нетрудно видеть, что выражение (4.2) представляет собой интеграл Шоке по мере доверия $Bel(A)=\sum_{B\subseteq A}\bar{m}(B)$ от функции $V_k=(v_{k,1},...,v_{k,2n})$.

Пример 4.1. Предположим, что в выборах участвовали три партии A, B и C, которые до выборов обозначили свою позицию по 4-м вопросам. Ответы на вопросы представлены в Таблице 4.1.

Таблица 4.1 – Позиции трех партий по 4-м вопросам

	A	B	C
1	1	1	0
2	0	1	1
3	1	0	0
4	0	0	0
	50%	40%	10%

На выборах эти партии набрали 50%, 40% и 10% голосов соответственно. Какие группы вопросов следует считать более значимыми? Предположим, что нас интересуют только пары вопросов. Тогда более значимыми будут те пары вопросов, на которые партии A и B, которые набрали наибольшее число голосов, ответили примерно одинаково, а ответы партии C наоборот, сильно отличались от ответов партий A и B. Нетрудно видеть, что наименьшие расхождение в ответах у партий A и B были на пары вопросов $\{1,4\}$ (0 расхождений), $\{1,2\}$, $\{1,3\}$, $\{2,4\}$, $\{3,4\}$ (1 расхождение). Но из этих пяти пар опросов наибольшее расхождение в ответах у партий A и B с партией C будет по парам вопросов $\{1,2\}$ и $\{1,3\}$ (по 3 расхождения). Эти пары и следует считать наиболее значимыми.

Пусть $\{A_j\}_{j=1}^{3^n-1}$ – некоторое упорядочение всех подмножеств из $2^{\overline{N_n}}$, $\mathbf{m}=(m_j)$, $m_j=\bar{m}(A_j)$, $c_{k,j}=\min\limits_{i\in A_j}\{v_{k,i}\}$, $j=1,...,3^n-1$, $k=1,...,r$, $\mathbf{c}_k$ – k-я вектор-строка матрицы $\mathbf{C}=(c_{kj})$. Тогда $C(V_k)=\mathbf{c}_k\mathbf{m}$ и верно утверждение.

Утверждение 4.1. Решение задачи минимизации функционала R при условиях (4.3) равносильно решению задачи

$$\begin{cases} \mathbf{m}^T\mathbf{Q}\mathbf{m} \to \min, \\ m_j \geq 0 \ \ \forall j, \\ \sum_j m_j = 1, \end{cases} \tag{4.4}$$

где $\mathbf{Q} = \sum_{k=1}^{r-1} (p_{k+1}\mathbf{c}_k - p_k\mathbf{c}_{k+1})^T (p_{k+1}\mathbf{c}_k - p_k\mathbf{c}_{k+1})$.

Заметим, что если для некоторого подмножества $A \in 2^{\overline{N_n}}$ $\min\limits_k \min\limits_{i\in A}\{v_{k,i}\} = 0$, то переменную, соответствующую этому множеству A, можно исключить из вектора $\mathbf{m}$. Поэтому реальная размерность векторов $\mathbf{m}$ намного меньше, чем 3^n.

Поскольку $\operatorname{rank}\mathbf{Q} \leq r \ll 3^n$, то задача (4.4) может иметь множество решений. Для решения такой некорректной задачи вместо функционала $\mathbf{m}^T\mathbf{Q}\mathbf{m}$ в (4.4) рассмотрим его регуляризацию $\mathbf{m}^T\mathbf{Q}\mathbf{m} - \alpha\|\mathbf{m}\|^2 = \mathbf{m}^T(\mathbf{Q} - \alpha)\mathbf{m}$, где $\alpha > 0$ – параметр регуляризации. Регуляризация больше «штрафует» те распределения $\mathbf{m}$, которые ближе к равномерному распределению.

Заметим, что минимальное значение критерия R может быть равным нулю, если система $p_{k+1}C(V_k) = p_kC(V_{k+1})$, $k = 1,\ldots,r$ имеет решение. Из (4.2) следует, что эта система равносильная линейной системе

$$\sum_{A\in 2^{\overline{N_n}}} \bar{m}(A)b_k(A) = 0,\ k = 1,\ldots,r,\ \sum_{A\in 2^{\overline{N_n}}} \bar{m}(A) = 1,\ \bar{m}(A) \geq 0 \text{ для всех } A \in 2^{\overline{N_n}},$$

где $b_k(A) = p_{k+1}\min\limits_{i\in A}\{v_{k,i}\} - p_k\min\limits_{i\in A}\{v_{k+1,i}\}$, $k = 1,\ldots,r$, $A \in 2^{\overline{N_n}}$.

После нахождения функции масс $m: 2^{N_n} \to [0,1]$ найдем функцию доверия $Bel(A) = \sum_{B\subseteq A} m(B)$, которая характеризует степень доверия к тому, что множество вопросов является значимым для голосования. Вклад отдельных вопросов в общую значимость оценим с помощью значений Шепли.

Политическую неоднородность общества оценим с помощью меры внутреннего конфликта тела свидетельств $F=(A\ ,m)$, определенного на множестве всех вопросов N_n. Если для различных электоральных групп значимыми являются совершенно разные непересекающиеся (или «слабопересекающиеся») группы предвыборных вопросов, то это говорит о большой неоднородности общества. В качестве меры внутреннего конфликта будем использовать величину $K_{\text{int}}(F)=\min_i Bel(N_n\setminus\{i\})$ (см. пункт 4.1.2).

4.1.3.2 Оценка влияния позиций партий на голосование

Рассмотрим проекции свидетельства значимости позиций для результатов голосования на профили отдельных партий.

Будем говорить, что подмножество $A\in 2^{\overline{N_n}}$ содержится в профиле k-й партии V_k (и обозначать это $A\subseteq V_k$), если $\forall i\in A: v_{k,i}=1$. Пусть $\mathcal{A}$ – множество всех фокальных элементов, найденных в предыдущей задаче, $\mathcal{A}_k=\{A\in\mathcal{A}\ : A\subseteq V_k\}$ – проекция фокальных элементов на профиль k-й партии и $\bar{m}_k(A)=\bar{m}(A)$, если $A\in\mathcal{A}_k$. Величина $\bar{m}_k(A)$ характеризует электоральную значимость для k-й партии позиции по множеству вопросов $A\in\mathcal{A}_k$. Тогда величина $M_k=\sum_{A\in\mathcal{A}_k}\bar{m}_k(A)$ характеризует степень значимости позиции k-й партии на результат голосования за нее. Нетрудно видеть, что $M_k\in[0,1]$ и если $M_k=0$, то позиция партии по списку вопросов незначима для выборов. Если же $M_k=1$, то позиция партии по всем вопросам значима для победы на выборах.

4.1.3.3 Оценки конфликтности позиций партий по значимым вопросам

При принятии решений о создании коалиций или других совместных действий существенное значение приобретает анализ близости позиций партий по ключевым вопросам. В рамках теории функций доверия это можно сделать, проанализировав конфликтность между свидетельствами электоральной значимости позиций партий по множеству вопросов.

Для этого по функциям множеств $\{\bar{m}_k\}$ сформируем новые тела свидетельств $\{F_k\}$, $F_k = (A_k, m'_k)$, $m'_k = \bar{m}_k / M_k$, если $M_k \neq 0$. Если же $M_k = 0$, то это говорит о незначимости позиции партии на результат выборов. Следовательно, в этом случае позиция k-й партии будет неконфликтной с позициями других партий в рассматриваемом смысле (хотя, конечно, может быть конфликтной относительно других мер).

Для оценивания степени конфликтности позиций партий будем использовать меру внешнего конфликта (4.1). Под конфликтностью (несогласованностью) позиций между i-й и j-й партиями будем понимать величину $K_{i,j} = K(F_i, F_j)$, $K_{i,j} \in [0,1]$. Значение $K_{i,j} = 1$ означает, что позиции i-й и j-й партий по всем вопросам и их различным значимым для голосования комбинациям не пересекаются. А значение $K_{i,j} = 0$ означает, что у i-й и j-й партий нет значимых для голосования непересекающихся позиций.

4.1.4 Описание данных

В статье использовались данные, полученные с сайта The German Federal Agency for Civic Education [219] о позициях 28 партий (или блоков) Германии по актуальным вопросам перед выборами в Бундестаг 2013г. Всего в списке – 38 вопросов, на каждый из которых каждая партия (или блок) должна ответить «да» или «нет». Кроме того, в исследовании использовались результаты выборов 2013 года, взятые с официального сайта немецкого парламента [233].

Для наглядности и уменьшения вычислительной сложности были отобраны только те 8 вопросов, за которые население отдало наибольшее число голосов. Это, например, вопрос #22 о повышении пенсионного возраста (за включение этого вопроса проголосовало 3 490 000 человек), вопрос #18 о системе государственного финансирования студентов (за него проголосовало 2 060 000 человек), вопрос #2 о субсидировании семей с детьми, не посещающих учебные заведения (за него проголосовало 1 010 000 человек) и т.д.
Кроме того, были проанализированы результаты ответов на вопросы для тех 6 партий (или блоков), которые набрали больше 3% голосов на парламентских выборах. Это блок Христианско-демократического союза и Христианско-социального союза Баварии (CDU/CSU), Социал-демократическая партия Германии (SPD), «Левые» (DIE LINKE), «Союз 90/Зелёные» (GRUNE), Свободная демократическая партия (FDP), «Альтернатива для Германии» (AfD).

4.1.5 Результаты исследования

Результаты нахождения наиболее значимых для голосования групп вопросов.

Для множества из 8-ми вопросов и 6 партий (или блоков) было выявлено 1003 подмножеств вопросов A, для которых $\min_{k} \min_{i \in A} \{v_{k,i}\} > 0$. Это и есть реальная размерность векторов $\mathbf{m}$.

При нахождении наиболее значимых для голосования подмножеств из 8-ми вопросов методом, описанным в подпункте 4.1.3.1, было найдено 255 подмножеств вопросов с ненулевыми значениями функции масс (фокальные элементы). Группа наиболее значимых подмножеств состоит из 30 наборов вопросов со значимостью 0.0072. Во все подмножества с наибольшей значимостью вошли вопросы #18, #22 и #2. Значимость именно этих вопросов подтверждает и вектор Шепли. Значения Шепли будут равны 0.18, 0.15, 0.12 для вопросов #2, #8, #22 соответственно и примерно 0.11 для остальных 5-ти вопросов.

Оценивание политической неоднородности общества с помощью меры внутреннего конфликта свидетельства значимости вопросов.

Эта мера (см. подпункт 4.1.3.1) в нашем случае будет равна $K_{\text{int}}(F) = 0.264$. Это говорит об определенной политической однородности общества (хотя, конечно, мера носит относительный характер).

Результаты оценка значимости позиций отдельных партий для голосования, вычисленные по формуле $M_k = \sum_{A \in A_k} \bar{m}_k(A)$, $k = 1, \ldots, r$ (см. подпункт 4.1.3.2) представлены на графике рисунке 4.1. Там же представлен и график долей голосов избирателей за каждую из 6 партий. Видно, что графики значимостей и долей голосов сильно коррелируют друг с другом.

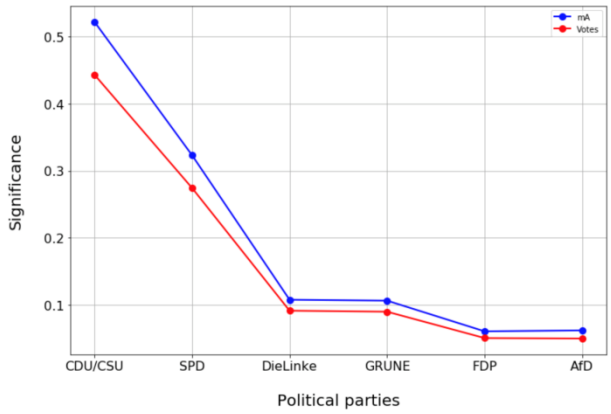


Рисунок 4.1 – Значимости позиций партий по множеству из 8 вопросов на результат голосования

Результаты оценки конфликтности позиций партий по значимым вопросам, вычисленные с помощью меры конфликта (см. подпункт 4.1.3.3), представлены на рисунке 4.2.

	CDU/CSU	SPD	DieLinke	GRUNE	FDP	AfD
CDU/CSU		0,35	0,59	0,42	0,14	0,14
SPD			0,12	0,16	0,40	0,20
DieLinke				0,13	0,32	0,48
GRUNE					0,20	0,36
FDP						0,18
AfD						

Рисунок 4.2 – Конфликтность позиций партий по значимым вопросам (чем краснее – тем конфликтнее, чем зеленее – тем менее конфликтны партии)

Заметим, что партия CDU/CSU имеет большую меру конфликта с партиями – ближайшими тремя преследователями, в особенности с партией DieLinke, наименьшая мера конфликта из этих трех партий получилась с партией SPD. А с партиями FDP и AfD по основным множествам вопросов значимым для голосования позиции партии лидера совпадают. Заметно, что оппозиция из партий SPD, DieLinke и GRUNE имеют схожие политические взгляды, что отражено в их значениях меры конфликта.

На предыдущих выборах в Бундестаг в 2009 году была сформирована коалиция из партий CDU/CSU и FDP. Модель подтверждает согласованность позиций данных партий. Партии AfD и CDU/CSU являются правыми партиями, что также подразумевает возможность их согласованности. Сопоставляя полученные результаты с реальными коалициями, можно сказать, что модель показала себя корректно.

4.1.6 Некоторые выводы

Методы теории неаддитивных мер (включая интегрирование по таким мерам), а также методы теории функций доверия (включая оценивание внутреннего и внешнего конфликта свидетельств) применены к решению ряда задач анализа значимости и конфликтности позиций политических партий. Исследование проведено по базе данных интернет-опросов партий в Германии перед выборами в Бундестаг в 2013г. и результатов этих выборов. В частности, показана возможность решения методами теории неаддитивных мер и функций доверия следующих задач:

– нахождение наиболее значимых для голосования групп вопросов путем минимизации расхождения между распределением голосов по партиям на выборах и априорным распределениям потенциальных голосов по партиям, вычисленным с помощью интеграла Шоке по данным интернет-опросов партий перед выборами;

– оценивание политической неоднородности общества с помощью меры внутреннего конфликта свидетельства значимости вопросов;

– оценивание значимости позиций отдельных партий для голосования с помощью построения проекций свидетельства значимости позиций для результатов голосования на профили отдельных партий;

– оценивание конфликтности позиций партий по значимым вопросам с помощью меры внешнего конфликта.

Результаты этого исследования отражены в [252, 253].

4.2 Исследования поляризованности и влияния в политических системах

4.2.1 Введение

Рассмотрение современных политических вопросов, связанных с Европой, представляется значимым в связи с важностью геополитической позиции данного региона. С этой точки зрения, Европа, несмотря на видимую сегодня со стороны однородность, в основном представляет собой сложного и неоднородного политического игрока, особенно учитывая тот факт, что в составе региона представлены различные по своему политическому устройству государства. Современные тенденции развития европейских стран отражают серьезные изменения, происходящие в их парламентах.

Парламент – высший представительный и законодательный орган государственной власти, основными функциями которого является отражение политических волеизъявлений граждан и формирование законодательной базы государства. В рамках стран с различной формой политического устройства можно наблюдать множество способов функционирования и отправления политики данного органа власти. Несомненным, однако, остается тот факт, что принятие решений здесь всегда происходит коллегиально. Тем не менее, степень влияния политических групп и коалиции внутри парламента всегда варь-

ируется и может быть измерена. Логически вытекающий вывод о том, что влиятельной является та фракция, которая набрала наибольшее число голосов на выборах, не всегда отражает действительную картину распределения влияния в парламенте.

Воплощение демократического режима и формирование государственного курса напрямую связаны с партийной ранжировкой и распределением мест в парламенте. Противостояние партий, которое выражается в поляризованности парламента, является отражением многообразия тех установок, которые выработались по ключевым политическим вопросам.

Поляризация может быть представлена разными дефинициями, однако в качестве основной чаще всего рассматривают определение, данное Эстэбаном и Рэем [254], в котором поляризация представлена как ситуация, когда при выборе какой-либо характеристики образуются кластеры, одинаковые внутри своего образования, но выступающие противоположностью по отношению к другому кластеру. Подход Эстэбана и Рэя связан с проблемой идентификации и отчуждения, когда наблюдается усиление связи между членами одного кластера, но возрастает отчуждение по отношению к другому.

Для того чтобы детально рассмотреть поляризованность и политическое влияние в парламентах, необходимо определить, от каких факторов нужно отталкиваться при изучении самого парламента. Роль его, прежде всего, напрямую зависима от типа политической системы, политического режима и формы правления, от формальных и неформальных политических правил игры, структуры, состава и стабильности данного политического института. В рамках исследования предлагается рассмотреть следующие показатели, влияющие на деятельность парламента: роль парламента в жизни государства, его взаимодействие с исполнительной ветвью власти, классификация существующей партийной системы и число эффективных партий, а также тип избирательной системы, по которой избираются члены парламента.

4.2.2 Политическое влияние в парламентах

Взаимодействие исполнительной и законодательной ветвей власти, а также описание задач, стоящих перед парламентом, представлено в [255, 256]. В этих работах выделено пять основных функций парламента: представительская, дискуссионная, экспертная, выработка и принятие решений, а также сдерживание исполнительной ветви власти и тенденций мажоритарности.

Описание партийных систем в политологии занимает отдельный важный блок. Однако для того, чтобы сузить рассмотрение вопроса с точки зрения его применимости к анализу влияния и поляризованности парламента, следует рассмотреть партийную классификацию Сиароффа [257]. По мнению Г. Голосова [258], он внес существенный вклад в развитие теории партийных систем, разработав наиболее полный и последовательный подход, который, несмотря на его изначальную применимость в отношении только послевоенных западноевропейских демократий, может быть расширен.

Показатель эффективного числа партий был разработан в [259] для оценки фрагментации партийных систем. Он показывает положение партии в партийной системе, ее относительный вес. Несмотря на простоту использования индекса, при подсчете он не отражает точные различия между системой с доминирующей партией и двухпартийной.

В ряде работ, описывающих партийную систему, так или иначе, затрагивается тема избирательных систем и их влияния на распределение влияния в парламенте. В политологии принято выделять два основных типа избирательных систем: мажоритарную и пропорциональную. Каждая из них представлена в нескольких вариациях. От выбранного вида избрания членов законодательного органа власти, способа подсчета и распределения голосов во многом зависит распределение влияния в парламенте.

Политическое влияние является отражением той силы, которую может продемонстрировать коалиция или партия в процессе принятия решений. Важность оценки данного показателя исходит из того факта, что работа парламента в демократических государствах нацелена на выражение интересов

населения, результатом ее служит формирование социально-экономической жизни государства. Для принятия политического решения необходим консенсус, который в парламенте достигается по-разному. Правила, принятые в парламенте, могут требовать достижения простого большинства, квалифицированного большинства, выражаемого 2/3 или 2/3+1 голосов.

Измерение политического влияния позволяет увидеть реальное распределение сил в парламенте, оценить направления работы законодательного органа власти, его эффективность. Еще одной целью оценки показателя может стать возможность проследить появление ситуаций политического застоя, выработка способов их преодоления.

Другим важным критерием, описывающим ситуацию, происходящую в парламентах, является поляризованность. В настоящее время изучение данного феномена привлекает все большее внимание со стороны научного сообщества. Прежде всего, рассматривают его с целью изучить то напряжение в социуме, которое существует в силу различных обстоятельств. Показатель отражает также кризисное состояние, характерное для общества в данный момент времени.

Оценивать полязированность и политическое влияние в парламенте невозможно без представления о том, какая политическая система представлена в государстве. Несомненно, что эта категория слишком широка для исследования в рамках узкого анализа работы парламента, однако возможно изучение отдельных ее аспектов, способных выступить точкой опоры при оценке данных. Сосредоточившись на изучении роли парламента в рамках политической системы и факторов, влияющих на его работу, можно сформулировать следующую проблему исследования. А именно: как изменения, произошедшие в законодательной и исполнительной ветвях власти европейскийх государств, влияют на поляризацию в парламенте и распределение политического влияния при принятиях решений.

Впервые с научной точки зрения к проблеме изучения влияния подошел Пенроуз [260], который, на примере оценки политического влияния стран-

представителей в Генеральной Ассамблее ООН, предположил, что оно пропорционально вероятности того, будет ли голос решающим в процессе принятия решений. Основываясь на расчетах данного автора, свою модель измерения влияния предложил Банцаф [261, 262, 263].

Описанию данного феномена посвящена монография Фельсенталя и Маховера [264], которые выразили его в теории априороного или взвешенного влияния. Отталкиваясь от идеи Л. Мартина [265] о ключевом игроке, рассматривалась возможность участников принятия решений в рамках голосования по принципу «да/нет» влиять на результат. В данной работе были описаны также парадоксы данного явления и новые исследования, учитывающие влияние ситуации неучастия субъекта в голосовании.

Первым исследователем, предложившим использовать индексы влияния в политической теории, стал также Пенроуз. Немногим позже появились исследования Райкера [266] о политическом влиянии. Основываясь на оценке работы французского парламента, он выдвинул идею о том, что межпартийные объединения позволяют увеличить влияние коалиций при принятии решений. Однако, столь упрощенное восприятие феномена подверглось критике.

В контексте развития теории коалиционных игр появился индекс Шепли [243], основанный на выявлении возможных коалиций. Совместными усилиями Шепли и Шубика [137], предпринятыми позднее, появляется индекс, позволяющий измерить влияние участников процесса принятия решений с помощью подсчета частоты события, при котором игрок становится ключевым. Однако, данный подход не учитывает предпочтения игроков, тем самым образуя серьезные погрешности результатов. Устранение представленного недостатка стало возможным благодаря отражению положения игрока в пространстве на основе его предпочтений, предложенном в концепции Шепли-Оуэна [267]. Таким образом, при парламентском голосовании возможно учитывать не только общепринятые правила, но и политические предпочтения участников.

В настоящее время появились исследования распределения политического влияния на основе обобщений индекса Шепли-Оуэна, предложенных

Ф. Алескеровым [268]. Данный подход основывается на «весе», приписываемом каждому ключевому игроку, на расчете согласованности их позиций – близости предпочтений в политическом пространстве. Вероятность вступления игрока в ту или иную коалицию зависит также от расстояния между идеальным положением игрока и «центром масс» игроков.

Индекс согласованности отображается как точка, занимаемая отдельным политическим игроком на политической карте. Описывая политическую карту, мы можем выразить другой, немаловажный показатель эффективности работы парламента, а именно поляризованность.

4.2.3 Почему важно изучать поляризацию

Отсутствие внутрипарламентских расколов позволяет формировать устойчивый законодательный орган власти. Показатель поляризации связан с его результативным функционированием и отсутствием долгих затяжных институциональных кризисов, а также с вопросом выживания государства в целом.

Ситуация ярко выраженной поляризации в рамках легислатуры создает проблему неустойчивости парламента и неэффективности его работы, поскольку нарушается возможность коммуникации и достижения консенсуса в процессе принятия политических решений. Возникновение политического конфликта может привести к образованию политического застоя, который, в свою очередь, может подорвать репутацию власти и стать причиной ее смещения. С другой стороны, в попытке предотвратить негативные последствия поляризации откроются благоприятные условия для формирования авторитарных тенденций.

Социальная поляризация может быть представлена разными дефинициями, однако в качестве основной чаще всего рассматривают определение, данное Эстэбаном и Рэем [254].

Их подход связан с проблемой идентификации и отчуждения, когда наблюдается усиление связи между членами одного кластера, но возрастает

отчуждение по отношению к другой. Большее стремление идентифицировать себя со своей группой усиливает поляризованность.

Ситуация ярко выраженной поляризации в рамках легислатуры создает проблему неустойчивости парламента и неэффективности его работы, поскольку нарушается возможность коммуникации и достижения консенсуса в процессе принятия политических решений. Возникновение политического конфликта может привести к образованию политического застоя, который, в свою очередь, может подорвать репутацию власти и стать причиной ее смещения. С другой стороны, в попытке предотвратить негативные последствия поляризации откроются благоприятные условия для формирования авторитарных тенденций.

Вопрос исключительно отрицательного влияния поляризации является дискуссионным. Существует несколько теоретических концепций, отражающих взаимосвязь демократического режима и уровня поляризованности [269]:

– поляризация ведет к демократическому упадку;

– демократический кризис ведет к повышению уровня поляризованности;

– поляризация способствуют демократическому развитию.

Степень партийной согласованности, в свою очередь, варьируется в зависимости от типа избирательной и партийной системы, политического режима, централизации и децентрализации политической власти. Формирование партийных коалиций и альянсов может зависеть от индивидуальных предпочтений депутатов, вытекающих из жизненного бэкграунда личности, его карьерных политических предпосылок, а также его партийной принадлежности. Важным является позиционирование партии в идеологическом и проправительственном/оппозиционном спектре.

4.2.4 Контекст изучения поляризованности

Чаще феномен поляризации рассматривается в экономическом контексте, когда наблюдается разница в доходах населения, в результате чего образуются противоположные группы. Тем не менее, большинство исследований нацелено на описание ситуаций, при которых поляризация является показателем конфликтности общества. Таковой является [270], в которой описывается этнический индекс поляризованности и возможность его использования для предотвращения насильственного развития событий. В [271] также указывается на свидетельство связи между выражением поляризованности и отражением уровня конфликтности в рамках исследуемой структуры. По мнению ученых, она позволяет также увидеть связь между разрозненностью и неравенством.

Введение медианы при оценке поляризованности было предложено в [272, 273]. Распределение кластеров в данном случае наглядно демонстрируется за счет отклонения от срединного показателя. Такой подход описывает биполязирацию и позволяет описать ситуацию в рамках одномерного пространства, когда за основу берется один критерий [274, 275, 276].

Сегодня возрастает внимание к рассмотрению полязированности в многомерном пространстве, однако пока можно наблюдать некоторые успехи в рамках небольшого числа работ, среди которых, например, можно отметить работу Джиглиарано и Мослером [277].

Долгое время политическая поляризованность измерялась по идеологической шкале. Изучение феномена реализовывалось при рассмотрении партийной поляризованности в идеологическом спектре в [278], где обосновывался вывод о центробежных тенденциях партий. Джованни Сартори определил поляризацию (в рамках изучения партийных систем), основываясь на идеологическом противостоянии политических акторов, с некоторыми ситуациями, где может наблюдатся феномен стремления партий к запросам «медианного избирателя» – концепции, обоснованной в [279]. Самым распростра-

ненным измерением являлась до сегодняшнего времени идеологическая дистанция между партиями, ее ключевыми игроками или избирателями по шкале слева направо. Однако партийное размежевание возможно не только в рамках идеологической шкалы. Формирование партийных расколов осуществляется посредством других механизмов.

Вместе с тем, как отмечают многие авторы, борьба на выборах часто не структурирована по классическому даунскому лево-правому идеологическому континууму. Несмотря на то, что идеологическая маркировка остается значимой как когнитивный ярлык, который позволяет проводить категоризацию политических предпочтений. Понятие «идеология» в данном случае используется в достаточно широком понимании. Отталкиваясь от идеи Энтони Даунса, приверженность той или иной идеологической платформе может быть связана с некоторым общим представлением о должном распределении общественных благ и политической власти [279]. Когда происходит поляризация, можно наблюдать апелляцию партийных групп и отдельных членов партий к символьным идеологическим атрибутикам, например, ярлыкам «правый», «центрист», «левый». Однако подобная идентификация не всегда связана с реальными сущностными отличиями, а служит отражением ценностной полярности. Например, в государстве правые партии могут «решительно поддерживать религиозные принципы в политике, подчеркивают значимость националистических символов, отстаивают большее государственное вмешательство в вопросы формирования образа жизни граждан и в целом ценностного порядка, стабильности. Левые партии являются космополитическими, поддерживают светское государство и защищают большие индивидуальные и гражданские свободы» [280].

Тем не менее, в настоящее время политическая борьба, в силу разных обстоятельств, не выстраивается в классическом даунсоновском лево-правом идеологическом континууме. Мы можем говорить о других альтернативных измерениях политических размежеваний: религиозность/светскость взглядов, идеи национализма/ космополитизма, поддержка традиционных/современных

ценностей, противоречия города/деревни, приверженность свободному рынку/командной системе и другие [269].

Измерение поляризованности обязательно связано с выбором одной или нескольких шкал, по показателям которых измеряется поляризация. Несмотря на большой список возможных идеологическому спектру альтернатив, выбор другого спектра вызывает немало сложностей. Многовариантность предложенных МкКоем позиций варьируется в зависимости от государственной повести дня и ключевых проблем, разрешение которых волнует политическую власть. Поиск, а более того, интерпретация ситуационных спектров может вызвать большие затруднения и не дать результата, который может быть использован для преодоления сложностей, вызванных поляризацией.

Сегодня все больший интерес вызывает вопрос лояльности и оппозиционности групп в рамках работы парламента. Наличие оппозиции – это основа существования демократического государства. Сартори [278] утверждал, что неответственная оппозиция, формируется именно при поляризованном плюрализме.

Однако, как однажды заметил Даль: «Сказать, где заканчивается правительство и начинается оппозиция – это упражнение в метафизике» [281]. Несмотря на проявившуюся тенденцию политической согласованности парламентских фракций во многих Европейских государствах, мы не можем говорить о единой модели, поскольку оппозиция в каждом отдельно взятом государстве действует по-разному в зависимости от партийной структуры, длительности существования, политических целей и стратегий борьбы за власть.

Оппозиционность можно определить через согласованность партии А с проводимой партией В политикой, однако такой подход будет неверен, поскольку не говорит об открытых действиях партийных групп. Даль, говоря об оппозиции, определил ее посредством следующих рассуждений: «Предположим, что А определяет направление некоторых аспектов управления конкретной политической системой в течение некоторого интервала… Предположим, что в течение некоторого интервала В не может определять курс управления

правительства, а также что B противостоит курсу правительства, заданному A. Тогда B – это то, что мы подразумеваем под «оппозицией» [281]. Оппозиционность фракции в парламенте отчетливо видна при работе результатами пленарных голосований.

4.2.5 Некоторые выводы

В основу исследования заложены как количественные, так и качественные методы. Количественные представлены расчетами индексов политического влияния и поляризованности, пространственным моделированием Пуля и Розенталя при построении политической карты парламентов выбранных государств. Качественные выражены в оценке политической системы на основе работ Шугарта и Кэри по изучению взаимоотношений исполнительной и законодательной ветвей власти, описаний партийных систем государств в соответствии с классификацией Сиароффа и эффективности партий Лаакло и Таагаперы. Планируется, в том числе, использование теоретико-игровых моделей, кластерного анализа, математической статистики, имитационного моделирования.

Сосредоточившись на изучении роли парламента в рамках политической системы и факторов, влияющих на его работу, можно поставить следующую задачу исследования. А именно: как изменения, произошедшие в законодательной и исполнительной ветвях власти государств, влияют на поляризацию в парламенте и распределение политического влияния при принятии решений.

Выборка стран, участвующих в исследовании, лимитирована доступностью и открытостью данных, необходимых для измерения используемых индексов. Результаты реализованного поиска ограничат круг государств, участвующих в анализе, не снизив релевантность проводимого исследования.

5 Исследования моделей в финансовой и банковской сферах

В этом разделе отчета приведены результаты исследований математических моделей и анализа данных в финансовой и банковской сферах. Эта работа проводилась в следующих направлениях:

– исследование влияния модельного риска на точность оценок величины риск-взвешенных активов;

– анализ структурной сложности данных по фондовому рынку S&P-500.

5.1 Исследование влияния модельного риска на точность оценок величины риск-взвешенных активов

5.1.1 Общая постановка задачи

Качество моделей, построенных в рамках ПВР (оценки кредитного риска на основе вутренних рейтингов), проверяется независимым подразделением валидации банка, а также национальным банковским регулятором (Банком России) в процессе официально проводимой первичной оценки (валидации) банковских рейтинговых систем, а после начала официального применения банком ПВР – в ходе специализированного мониторинга, последующего контроля и надзора. Применение банком моделей ПВР допустимо лишь при условии их достаточно высокого качества с точки зрения его ключевых характеристик (дискриминационной силы, точности, стабильности, устойчивости).

Предполагается, что модели ПВР, обладающие более высоким качеством, позволяют более точно оценивать кредитный риск по формулам соглашения «Базель II», в первую очередь, формулам для определения размеров взвешенных по риску активов (Risk-Weighted Assets, RWA) и ожидаемых потерь (Expected Losses, EL). Можно ожидать, что модели низкого качества могут существенно недооценивать величину RWA. При этом даже модели ПВР приемлемого качества, удовлетворяющие минимально допустимым требова-

ниям и прошедшие процедуру валидации, оценивают RWA с различной точностью, что порождает возможности ее значительной недооценки. Для наглядности рассмотрим пример.

Предположим, банк разработал и успешно защитил перед регулятором модель PD (вероятности дефолта) с достаточно высокой дискриминационной способностью (например, индекс Джини (Gini) = 70%). В процессе эксплуатации в связи с недостаточной стабильностью и/или неустойчивостью модели, ее повышенной чувствительностью к изменяющимся внутренним и/или внешним условиям дискриминационная способность снижается до уровня Gini = 50%, который тем не менее остается вполне приемлемым[5]. Однако такое ухудшение качества модели никак не учитывается при расчете величины RWA по формулам соглашения «Базель II».

Таким образом, отсутствие прямого учета качества моделей ПВР в формулах расчета RWA создает для банков возможности арбитража, когда последним выгоднее предоставить регулятору заведомо несовершенную модель, занижающую оценку кредитного риска и соответствующий необходимый капитал. По этой причине целесообразно изучить данный вопрос и постараться оценить степень недооценки кредитного риска моделями ПВР с пониженным (но все же достаточным) качеством. После такой оценки можно определить, какой должна быть система компенсации указанного недостатка и стимулирования банков к улучшению качества моделей ПВР. Независимо от решений регулятора о применении описываемого подхода, руководство, заинтересованное в получении объективных оценок кредитного риска отдельно взятого банка, использующего ПВР-модели, может самостоятельно применять данную методику в виде KPI для разработчиков и валидаторов моделей.

[5] Например, считается, что приемлемым (хорошим) качеством обладает такая модель, которая показывает Gini более 60–70% при разработке и не менее 50–60% при валидации в зависимости от типа заемщиков (юридические или физические лица) и от типа модели в случае физических лиц (заявочная или поведенческая).

5.1.2 Обзор литературы

С одной стороны, свойства статистических моделей зависят от количества, состава и качества данных, лежащих в основе их построения и применения. В частности, при малой наблюдаемой выборке и/или малом числе дефолтов (низкодефолтные портфели — Low-Default Portfolio, LDP) оценки компонентов кредитного риска могут оказаться ненадежными, иметь высокую дисперсию, что повышает возможные погрешности оценки. Например, в [282] обосновано, что в анализируемой выборке из кредитного портфеля необходимо наличие не менее 20 дефолтных наблюдений независимо от общего количества наблюдений. Для низкодефолтных портфелей авторы предлагают вводить регуляторную надбавку к PD. Например, для количества дефолтов менее 20 при уровне значимости 75% и корреляции активов R = 12% для полученной банком оценки PD = 1% рекомендуется регуляторная надбавка в 2%, т.е. в расчете RWA предлагается использовать регуляторное значение PD = 3%.

С другой стороны, адекватность оценки величины кредитного риска определяется методикой моделирования. Важно отметить, что в рамках ПВР расчет RWA осуществляется в два этапа. Вначале по внутренним методикам и моделям банк оценивает значения компонент кредитного риска (PD, LGD и иных). Затем они подставляются в качестве факторов в фиксированные, заданные регулятором формулы. Ключевой из них является формула, реализующая модель [283] и основанная на ряде важных и небесспорных теоретических предпосылок, в частности, на следующих допущениях.

1) Нормальность закона распределения системного и идиосинкратического факторов кредитного риска. В [284] обосновывается целесообразность корректировки данного допущения (перехода с нормального на логнормальное распределение) и необходимость увеличения величины RWA примерно на 100%.

2) Взаимная некоррелированность компонент кредитного риска (PD, LGD, EAD). Исследование данной проблемы на российских данных представлено в [285].

3) Бесконечная делимость портфеля ссуд, т.е. предположение о стремлении количества ссуд в портфеле к бесконечности. Калибровка RWA с учетом данного предположения рассматривается в [286].

В текущей работе мы не исследуем качество самой надзорной формулы (включая ее предпосылки о нормальности, использованные в [283]), поскольку они заслуживают отдельного исследования. Цель исследования – рассмотреть влияние качества внутрибанковских моделей PD на искомую величину кредитного риска, определить возможную «цену» ошибок, допускаемых банком при оценке вероятности дефолта. В отличие от [282] в данном исследовании предлагается делать надбавку не к PD, а к итоговой величине RWA.

5.1.3 Описание методологии исследования

Для решения поставленной задачи рассмотрим крайние ситуации, характеризующие два свойства моделей ПВР в отношении их ключевых статических свойств (дискриминационной способности и точности) на примере моделей PD. Дискриминационная способность таких моделей зависит от знания банком фактического закона распределения интегрального показателя кредитного качества (скорингового балла) оцениваемых КТ. При этом точность модели определяется верным выбором порога отсечения для этого показателя на данном распределении.

Начнем с дискриминационной способности. Совершенную дискриминационную способность модели (совершенное ранжирование) будем условно ассоциировать с тем, что банку известен некий совершенный интегральный фактор Z (скоринговый балл), определяющий качество КТ единственно и абсолютно верно, на который банк ориентируется при оценке его PD (рисунок 5.1, сплошная линия с маркером квадрата). Полное отсутствие дискриминационной способности у модели PD (случайное ранжирование) может быть условно

соотнесено с тем, что банк ориентируется на некий фактор Y, абсолютно не связанный с качеством КТ (и соответственно с фактором Z) (рис. 1, пунктирная линия). Кроме того, банк может ранжировать КТ в порядке, обратном совершенному (обратное ранжирование), используя фактор –Z (рисунок 5.1, сплошная линия с маркером треугольника). В этом случае модель PD может прогнозировать точное количество дефолтов, однако дефолты предписывать «хорошим» заемщикам, т.е. действовать наоборот. Дискриминационную способность моделей PD будем измерять показателем Джини (или отношением точности — Accuracy Ratio, AR), который наиболее употребим для этой цели в рамках ПВР [287].

Смоделируем для каждого КТ случайные значения указанных выше факторов Z и Y. Для упрощения без потери общности полагаем Z и Y некоррелированными случайными величинами со стандартным нормальным распределением: $Z, Y \sim N(0; 1)$, $\mathrm{corr}(Z, Y) = 0$.

На основании дискриминирующего фактора X, на который ориентируется банк, определим вероятность дефолта КТ следующим образом:

$$PD = (1 + \exp(-X))^{-1},$$

где $X = \alpha \times Z + \beta \times Y$, α и β — параметры, меняющиеся от 0 до 1. В общем случае $\beta = 1 - \alpha$ (кроме случая обратного ранжирования, для которого $\alpha = -1$ и $\beta = 0$).

Необходимо признать, что возможным ограничением настоящего исследования является использование в нем допущения о нормальности интегрального фактора Z (впрочем, использованного также и в модели [283]). Для уточнения полученных выводов целесообразно оценить размер надбавки при использовании иных распределений Z (равномерное, гамма-распределение и т.п.).

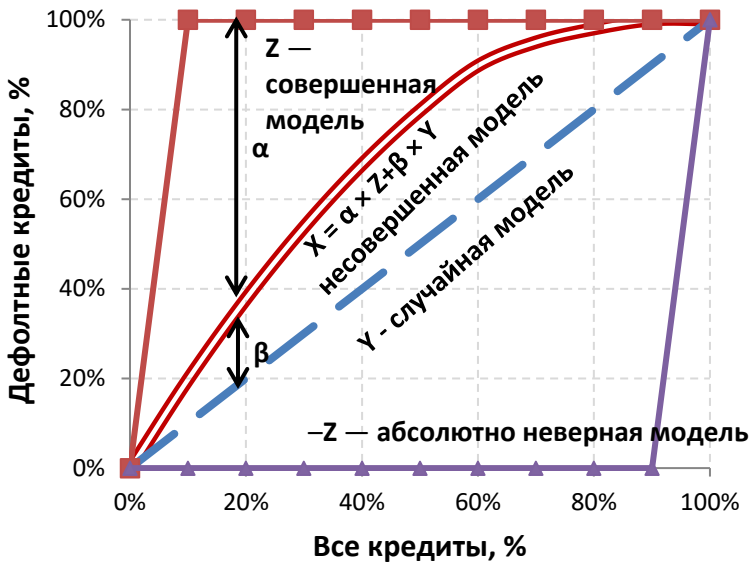


Рисунок 5.1 – Кривая Джини кумулятивного профиля точности, характеризующая дискриминационную способность модели PD

Точность модели будем ассоциировать со степенью совпадения прогнозного и фактически наблюдаемого количества дефолтов (рисунок 5.2), поэтому совершенная точность соответствует ситуации, когда для конкретной выборки прогнозируется количество плохих КТ, соответствующее их фактической (истинной) доле в этой выборке. Неточность модели может повлечь как недооценку этой фактической доли (что чаще интересует регулятора), так и его переоценку (что не приветствуется банком). Для цели текущего исследования мы будем измерять точность с помощью биномиального теста и теста Брайера (Brier Score), наиболее распространенных в статистических процедурах валидации внутрибанковских моделей ПВР [287].

В зависимости от целевого (запланированного) уровня дефолтов банк назначает необходимый уровень отсечения (квантиль распределения) для фактора, определяющего дискриминационную способность (Z или Y).

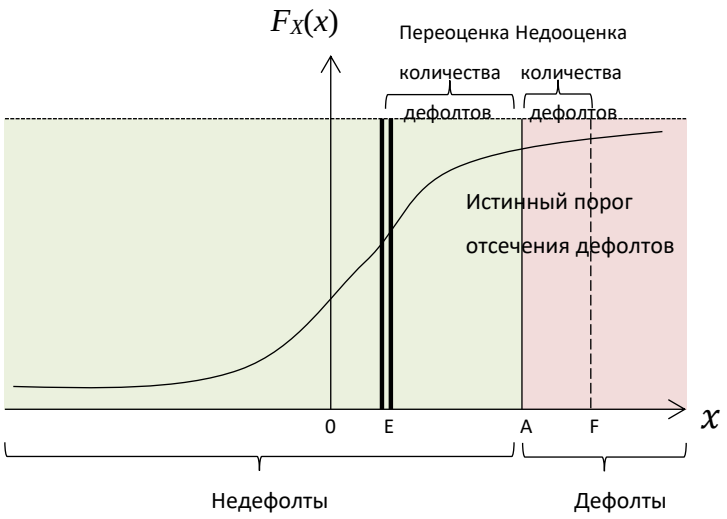


Рисунок 5.2 – Кривая распределения КТ по скоринговым баллам, применяемая для измерения точности модели PD

Пусть совершенная граница отсечения дефолтных КТ от недефолтных по выбранному дискриминирующему фактору X соответствует значению $T_{cut-off}$. КТ с фактором выше $T_{cut-off}$ через год окажутся в дефолте. Если граница отсечения по модели банка выше совершенной границы, то модель банка недооценивает количество дефолтов, которые произойдут через год, если ниже – переоценивает. При совпадении совершенной и модельной границы считаем, что модель абсолютно точна. Границу определяем как квантиль распределения фактора модели. Например, если квантиль берется равным 80%, то будет 20% дефолтов. Если банк по итогам разработки модели в качестве границы отсечения дефолтов выберет 90%-ный квантиль, значит, модель предполагает, что дефолтными станут только 10% вместо истинных 20% КТ.

Рассмотрим влияние качества модели PD на величину RWA. Для этого покажем проблемы, возникающие при формировании требований к капиталу под дефолтные и недефолтные активы.

В модели [283], использованной Базельским комитетом по банковскому надзору (БКБН), применяется концепция, согласно которой кредитные потери (или кредитный риск, его оценка) представляются в виде суммы ожидаемых (EL) и неожидаемых (Unexpected Losses, UL = RWA) кредитных потерь (рисунок 5.3). Под ожидаемые потери формируются резервы, под неожидаемые –

капитал в размере *RWA* × *CAR*, где *CAR* (Capital Adequacy Ratio) – норматив достаточности капитала [288].

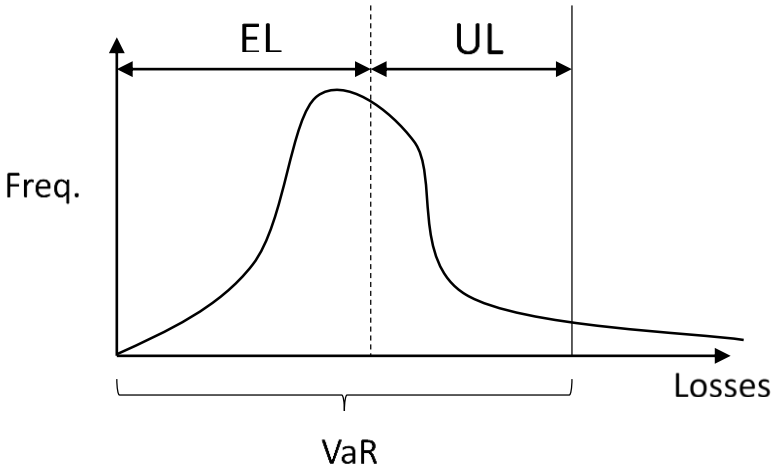


Рисунок 5.3 – Разложение потерь на ожидаемые и неожидаемые [288]

Между тем граница между ожидаемыми и неожидаемыми потерями условна. При этом изменение данной границы неоднозначно влияет на норматив достаточности капитала. Рассмотрим два крайних случая для суммарного кредитного риска 100 у.е. (*EL* + *UL* = 100) и величины капитала *K* = 200 у.е.

В первом случае положим *EL* = 99 и *UL* = 1, т.е.:

$$CAR_1 = \frac{K - EL}{12,5 \times UL} \times 100\% = \frac{200 - 99}{12,5 \times 1} \times 100\% = 16,08\%.$$

Во втором случае положим *EL* = 1 и *UL* = 99, т.е.:

$$CAR_2 = \frac{K - EL}{12,5 \times UL} \times 100\% = \frac{200 - 1}{12,5 \times 99} \times 100\% = 808\%.$$

Таким образом, при одинаковых начальных условиях смещение границы между *EL* и *UL* может привести к существенному изменению достаточности капитала. По этой причине в текущей работе мы рассматриваем суммарную величину кредитного риска (*EL* и *UL*). Ее отклонение от рассчитанной банком величины будем называть надбавкой к *RWA*. Исследование того, как оптимально учитывать штраф за модельный риск (в числителе, в EL, или в знаменателе, в RWA), выходит за рамки текущей работы.

Кроме того, согласно стандартам соглашения «Базель II» [289] для дефолтных активов кредитный риск оценивается (в долях от *RWA*) как разница между оценкой *LGD* и наилучшей оценкой ожидаемых потерь (не снижающейся ниже нуля). Тогда:

$$CAR = (K - EL) / RWA_{недефолт}.$$

В отличие от стандартов соглашения «Базель II», в российском регулировании (см. [290]) кредитный риск для дефолтных активов оценивается различно при использовании базового (БПВР) и продвинутого (ППВР) подходов. Для ППВР сохранен подход, представленный в соглашении «Базель II». Для БПВР по дефолтным активам оценка *RWA* определена как разница между 100% и ставкой фактически сформированных банком резервов.

$$CAR = (K - EL) / (RWA_{недефолт} + RWA_{дефолт}).$$

Это сделано для предотвращения манипуляций банков с капиталом. Иначе банкам было бы выгодно выводить актив в дефолт как можно раньше, тем самым обнуляя *RWA* без роста резервов, что приводило бы к искусственному повышению норматива достаточности капитала.

Следовательно, для КТ с фактором выше уровня отсечения назначается дефолт (PD = 100%) и величина взвешенных по риску активов определяется по формуле:

$$RWA_{дефолт} = (1 - PD \times LGD) \times EAD = (1 - 100\% \times 45\%) \times EAD = 55\% \times EAD.$$

При расчете ожидаемых потерь в целях упрощения расчетов полагаем, что для КТ LGD = 45%.

Получается, что совокупная оценка риска по российскому регулированию выше оценки риска по соглашению «Базель II» за счет добавления *RWA* по дефолтным активам, поэтому далее надбавки рассчитываются на основе российского регулирования.

5.1.4 Причина недооценки кредитного риска

При несовершенном (ниже 100%-ного) уровне качества модели PD недооценка кредитного риска возникает вследствие того, что банк прогнозирует выход в дефолт одного состава плохих КТ, тогда как фактически дефолтным может стать другой состав КТ в ином количестве (рисунок 5.4).

Банку необходимо держать капитал для всех КТ, по которым в начале анализируемого одногодичного периода прогнозировался дефолт: как фактически ставших (ИД), так и не ставших (ЛН) дефолтами. Понимая несовершенство применяемой модели, можно было бы предложить банку держать капитал без учета ложных недефолтов (ЛН). Однако КТ, не ставшие дефолтными до расчетной даты, вполне могут успеть выйти в дефолт до конца периода прогноза.

		Фактическая дефолтность КТ	
		Дефолтные (1)	Недефолтные (0)
Прогнозная дефолтность КТ	Дефолтные (1)	Истинные дефолты (ИД)	Ложные недефолты (ЛН)
	Недефолтные (0)	Ложные дефолты (ЛД)	Истинные недефолты (ИН)

Рисунок 5.4 – Матрица сопряженности бинарной классификации КТ по признаку дефолтности

Вместе с тем запас капитала требуется также и для тех КТ, которые фактически станут дефолтными вопреки прогнозным ожиданиям (ложные дефолты, ЛД). В результате такой реализации модельного риска растут ожидаемые потребности в капитале. Хотя неиспользованный капитал под ложные недефолты (ЛН) может быть использован для полной или частичной компенсации потерь по ложным дефолтам (ЛД), в текущей работе такая компенсация не учитывается, а предлагаемая надбавка (Н) является оценкой сверху для дополнительных требований к капиталу под модельный риск.

5.1.5 Пример количественной оценки недоучета кредитного риска

Рассмотрим девять возможных ситуаций для трех вариантов дискриминационной способности и трех вариантов точности. Совокупный кредитный риск рассчитываем как сумму *EL* и *RWA*. Все величины *EAD* КТ рассматриваемого портфеля для упрощения полагаем равными 100 у.е. Расчеты проведены для 1000 КТ.

Пусть через год в дефолт уходит 20% КТ (*DR* = 20%). Будем считать, что модель обладает избыточной точностью, если прогнозирует выход в дефолт более 20% КТ (возьмем для примера 35%), и заниженной точностью, если прогнозирует менее 20% КТ (возьмем для примера 5%). Модель обладает совершенной точностью, если прогнозирует ровно 20% дефолтов.

В Таблице 5.1 показано распределение *RWA* по вариантам соответствия прогнозных и фактических дефолтов в анализируемом кредитном портфеле.

Итоговые оценки *RWA* для различных сочетаний дискриминационной способности и точности модели PD представлены в Таблице 5.2.

Таблица 5.1 – Распределение *RWA* по соответствию прогнозной и фактической дефолтности для случаев реальной и совершенной моделей PD, у.е.

Источник прогноза	Прогнозная дефолтность КТ	Фактическая дефолтность КТ		Общий итог
		D (дефолтные)	ND (недефолтные)	
Реальная модель	D	8	28	36
	ND	4	18	22
Итого для реальной модели		12	46	59
Совершенная модель	D	8	10	18
	ND	12	18	31
Итого для совершенной модели		20	29	49
Максимальные значения (реальная и совершенная модели)	D	8	28	36
	ND	12	18	31
Требуемый показатель *RWA*				67

Таблица 5.2 – Итоговые оценки RWA для различных сочетаний дискриминационной способности и точности модели PD, у.е.

Фактор	Дискриминационная сила (индекс Джини, %)	Точность		
		Заниженная (5%)	Совершенная (20%)	Избыточная (35%)
–*Z*	Обратная (–100)	40	49	57
Y	Нейтральная (0)	41	49	59
Z	Совершенная (100)	41	49	59

При применении модели с совершенным ранжированием (*Gini* = 100%) и совершенной точностью (*DR* = 20%) требования к капиталу составят 49 у.е. Рассмотрим пример расчета недооценки *RWA* применительно к ситуации с избыточной долей дефолтов (*DR* = 35%) при полном отсутствии дискриминации, т.е. случайном ранжировании (*Gini* = 0%). Тогда требования к капиталу составят 59 у.е., т.е. больше 49 у.е. В этом случае банк недооценивает капитал по тем КТ, которые фактически станут дефолтными, т.к. оценивает их как 4 у.е. вместо необходимых 12 у.е. В результате вместо 59 у.е. потребуется 59 + (12 – 4) = 67 у.е. капитала. Результаты расчетов сведены в Таблице 5.3, отражающей величину капитала, необходимого банку вследствие того, что модель PD обладает низким качеством.

Таблица 5.3 – Необходимая величина капитала на покрытие кредитного риска, учитывающая недостаточное качество модели PD, у.е.

Фактор	Дискриминационная сила (индекс Джини, %)	Точность		
		Заниженная (5%)	Совершенная (20%)	Избыточная (35%)
–*Z*	Обратная (–100)	55	64	72
Y	Нейтральная (0)	53	60	67
Z	Совершенная (100)	49	49	58

Полученные величины *RWA* пересчитаем в процентные значения надбавки. Важно, что надбавка необходима не к истинной (49 у.е.), а к спрогнозированной банком (59 у.е.) величине *RWA*. В данном примере к *RWA* требуется надбавка 14% (Таблица 5.4):

$$(67 / 59) – 1 = +13{,}6\% \sim +14\%.$$

Таблица 5.4 – Необходимая величина надбавки к капиталу на покрытие кредитного риска, учитывающая недостаточное качество модели PD, %

Фактор	Дискриминационная сила (индекс Джини, %)	Точность		
		Заниженная (5%)	Совершенная (20%)	Избыточная (35%)
–*Z*	Обратная (–100)	(55 / 40) – 1 = +37,5	(64 / 49) – 1 = +30,6	(72 / 57) – 1 = +26,3
Y	Нейтральная (0)	(53 / 41) – 1 = +29,3	(60 / 49) – 1 = +22,4	(67 / 59) – 1 = +13,6
Z	Совершенная (100)	(49 / 41) – 1 = +19,5	(49 / 49) – 1 = +0	(58 / 59) – 1 = –1,7

Хотя есть такие комбинации характеристик модели, при которых сформированных требований к капиталу может быть больше, чем реально требуется, и размер надбавки получается отрицательным (например, для модели с совершенной дискриминационной силой и избыточной точностью надбавка составляет –1,7%), в дальнейшем под надбавкой будем понимать только неотрицательное изменение *RWA*.

5.1.6 Калибровка надбавки

Выше продемонстрирован подход для идеальных ситуаций с крайними проявлениями исследуемых свойств (полная дискриминация, абсолютная точность и противоположные случаи), которые фактически не встречаются на практике. По этой причине интересно рассмотреть более широкий спектр комбинаций значений точности и дискриминационной способности. Для этого проанализируем изменение параметра α от 0 до 1 с шагом 0,1, т.е. от случайной модели (*Gini* = 0%) до совершенно дискриминирующей модели (*Gini* = 100%). Получаем 10 комбинаций с точки зрения дискриминационной способности (Таблица 5.5). Модели с обратной (по отношению к верной) дискриминацией не рассмотрены, хотя они и вероятны с практической точки зрения. Уровень точности приведен по строкам и не влияет на расчет индекса Джини как индикатора дискриминирующей способности.

Таблица 5.5 – Размеры надбавки H (в процентах от RWA) для доли дефолтов DR = 10% (количество заемщиков N = 1000, количество дефолтов D = 100)

Отклонение фактической DR от пороговой по биномиальному тесту, %	Тест Брайера	Доля дефолтов по модели в общем количестве заемщиков, %	Показатели качества модели (индекс Джини, %)										
			100	99	97	93	84	73	58	42	26	14	3
82	0,109	20	0	0	1	2	4	7	8	9	10	11	11
63	0,104	18	0	0	1	3	5	7	8	9	11	11	12
40	0,101	16	0	0	1	3	6	8	10	11	11	12	12
22	0,101	14	0	0	2	4	7	9	10	11	12	13	13
2	0,101	12	0	1	4	6	9	10	11	12	13	13	14

Продолжение таблицы 5.5

–17	0,103	10	0 (СМ)	3	5	8	10	11	13	13	14	14	14
–38	0,105	8	4	5	8	11	12	13	14	15	15	15	15
–54	0,107	6	6	8	10	13	13	14 (А)	15	15	15	16	16
–70	0,110	4	9	10	12	13	14	15	16	16	16	17	17
–84	0,111	2	12	13	14	15	16	16	16	16	17	17	17

Примечание 5.1. Ограничения биномиального теста при валидации рейтинговых моделей и необходимость сравнения распределений прогнозных значений PD и фактических DR из-за бимодальности распределения последнего при коррелированных дефолтах подробно рассмотрены в [291].

Примечание 5.2. Использованы следующие обозначения: н/р — не рассчитывалось; СМ — совершенная модель; А — пример надбавки (14% к RWA) при отклонении от совершенной модели при уровне дискриминирующей способности Gini = 73% и заниженной точности модели в 6% при истинной доле в 10%.

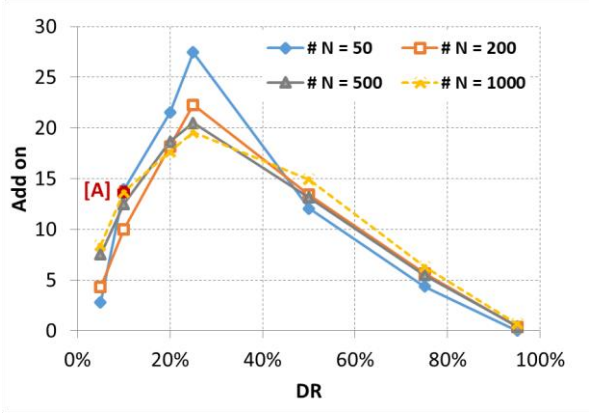
В левой части Таблицы 5.5 представлены два подхода к оценке точности: определение отклонения от порога биномиального теста без корреляции и тест Брайера. Соответствующие значения были рассчитаны для каждой из возможных комбинаций показателей точности и дискриминирующей способности.

Совершенная (по дискриминирующей способности и точности) модель, примененная к 1000 КТ с 100 истинными дефолтами (DR = 100 / 1000 = 10%), обозначена как СМ. Для нее надбавка к RWA составляет 0%. Недооценка DR на 4 пп. (указано как сдвиг вниз на две строки) и снижение $Gini$ до 73% (сдвиг на пять столбцов вправо) отражается как смещение из ячейки СМ в ячейку А, что соответствует 14%-ной надбавке к RWA.

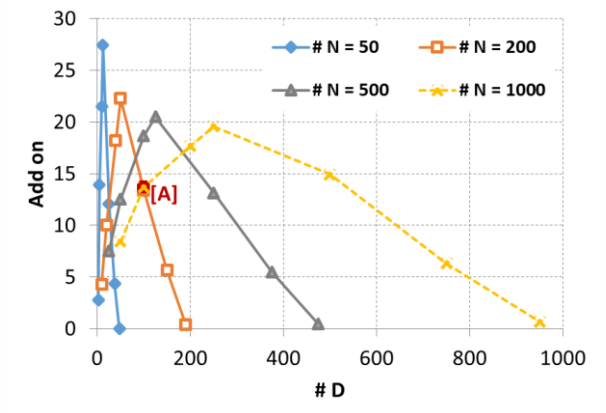
Надбавка (H) указана в процентах от первоначальной величины RWA_O, т.е. итоговая величина RWA определяется как:

$$RWA = \left(1 + \frac{H}{100}\right) RWA_O .$$

Например, при применении к портфелю из 1000 КТ модели PD с дискриминационной способностью *Gini* = 73% при целевой доле дефолтов 6% и их фактической доле 10% требуемая надбавка к RWA составляет 14% (см. Таблицу 5.5).



а) DR



б) D

Рисунок 5.5 – Калибровка величины надбавки к RWA для фиксированного количества наблюдений

Калибровка величины надбавки к RWA показана на рисунке 5.5. Оценка надбавки приводилась при фиксации одного из трех параметров: число наблюдений N, число дефолтов D и доля дефолтов в выборке DR. Чувствительность модели к изменению ее точности исследована для нескольких уровней доли дефолтов в выборке (DR = 5%, 10%, 15%, 20%).

На каждом графике для наглядности обозначена точка A из Таблицы 5.5. Таким образом, каждая точка на графике — это точка смещения на –4 пп. по точности и до 73% по *Gini* (т.е. линии строятся на основе данных из разных таблиц).

Визуальный анализ полученных графиков позволяет сформулировать следующие общие выводы.

1. Надбавка максимальна при *DR* в диапазоне от 25% до 30% (при таком *PD* вычисленная по формуле О. Васичека величина *RWA* максимальна). При более высоком *DR* размер надбавки снижается вследствие уменьшения вели-

чины непредвиденных потерь и их перехода в ожидаемые потери, покрываемые резервами. При *DR* меньше, чем 25–30% надбавка снижается (см. рисунок 5.5а).

2. При малом числе дефолтов или наблюдений надбавка требуется практически всегда, и она выше, чем при большем числе дефолтов или большем числе наблюдений (со снижением числа наблюдений — при смещении влево (см. рисунок 5.5б)).

3. С ростом числа дефолтов надбавка растет при $DR < 25\%$ и снижается при $DR > 25\%$ (см. рисунок 5.5б).

5.1.7 Основные выводы

В настоящем исследовании получена оценка сверху «цены» ошибки модели ПВР при ее минимально приемлемом (но все же достаточном) качестве за счет реализации модельного риска. Мы рассмотрели, как она может учитываться в текущих подходах к регулированию расчета *RWA* в рамках реализации ПВР. Подведем итоги.

Во-первых, ошибка в оценке *RWA* может стать причиной того, что два банка, оценивая при прочих равных условиях одного и того же заемщика, могут получить разные требования к капиталу на покрытие его кредитного риска. Это отметил БКБН [292] при наблюдении существенного разброса от –2 до +2 пп. в оценках требований к капиталу (при среднем уровне норматива 8–10 пп.). По этой причине, возможно, что выявленная БКБН проблема может быть следствием того, что в формуле расчета *RWA* по методологии соглашения «Базель II» не учтено влияние качества моделей, с помощью которых на практике осуществляется оценка компонентов кредитного риска. Мы полагаем, что введение такого учета (например, с помощью предлагаемых в настоящей работе надбавок) может позволить регуляторам стимулировать банки к улучшению своих внутренних моделей ПВР. В итоге это может повысить сопоставимость оценок кредитного риска.

Во-вторых, текущая практика регулирования в сфере ПВР уже включает надбавку за неточность модели – множитель 1,06 в формуле *RWA* [289], что, по сути, и есть надбавка в 6%. По нашим оценкам, надбавка в 6% эквивалентна ситуации наличия точной, но несовершенно дискриминирующей модели с *Gini* ≈ 70%. Не исключено, что эксперты БКБН проводили аналогичную калибровку, и потому в формулу *RWA* заложена указанная надбавка, соответствующая полученной в текущем исследовании для *DR* = 5%, *N* = 1000. Эта надбавка вводилась на переходный период, после которого ожидалось, что модели усовершенствуются и надбавку можно будет удалить. С учетом наблюдающегося в настоящее время недостаточно высокого качества моделей ПВР запланированная в финальной версии соглашения «Базель III» [293] отмена надбавки в 1,06 представляется необоснованной. Более того, полагаем, что зачастую необходима даже более высокая надбавка (более 6%, предусмотренных в настоящее время).

5.2 Анализ структурной сложности данных по фондовому рынку S&P-500

5.2.1 Введение

Представленный отчёт посвящён дальнейшему развитию предложенного ранее в публикациях [294, 295, 296, 297, 298, 299] исследования структуры графов. На его основе были введены численные индексы, характеризующие сложность графа, разработаны эффективные алгоритмы их определения и указаны перспективные направления применения разработанного инструментария для анализа различных систем, моделируемых графами.

Основные отличия между данной и предшествующими работами могут быть коротко выражены следующим образом.

1. В отличие от введённого ранее энтропийного индекса, рассмотрен более простой и наглядный индекс, равный числу различных найденных дихотомий графа, без вычисления энтропии соответствующего разбиения множества всех построенных дихотомий на группы совпадающих.

2. Впервые были рассмотрены ежедневные данные фондового рынка S&P-500 на протяжении почти 30 лет – с начала 1990 года до 30.09.2019. Речь идёт не только об обобщающих характеристиках, но и о ежедневных курсах акций всех компаний, входящих в S&P-500 с учётом изменений состава этого важного сегмента фондового рынка. Построенные наборы чисел для 10500 последовательных дней являются практически полными исходными данными для анализа процессов, проходящих на фондовом рынке. Данные для каждого дня вычисляются на основе только предшествующей информации, относящейся к достаточно коротким промежуткам (15-20 непосредственно предшествующих дней).

Полученные схемы можно использовать для анализа и представления динамики самых разнообразных процессов, происходящих в системах, моделируемых графами.

5.2.2 Частотный алгоритм дихотомии

В настоящем отчете кратко излагается недавно предложенный новый алгоритм построения семейства дихотомий произвольного неориентированного простого графа, опубликованный в статье [299]. Поскольку графы других типов в работе не рассматриваются, будем далее пользоваться термином «граф» без дополнительных уточнений.

Дадим описание общей структуры частотного алгоритма дихотомии, далее для краткости обозначаемого ЧАД.

1) На предварительном этапе каждому ребру исходного графа сопоставляется число 1. Далее числа, сопоставляемые рёбрам графа, будут называться частотой соответствующего ребра. Положим $F_{\max} = 1$.

2) Как и в предыдущих публикациях [296, 297, 298, 299], сам ЧАД состоит в многократном последовательном выполнении основного этапа. В результате каждого выполнения этого этапа частоты некоторых рёбер увеличиваются на 1, а во всех остальных рёбрах частоты не изменяются. Опишем основной этап подробнее.

1) С помощью стандартного датчика равномерно распределённых случайных чисел выбираются две различные вершины графа.

2) Для двух выбранных вершин алгоритмом Дейкстры находится соединяющий их кратчайший путь. Длиной ребра является его текущая частота. Длина пути равна длине его самого длинного ребра, а не сумме длин всех его рёбер. Хорошо известно, что алгоритм Дейкстры применим и в подобных случаях, с единственным изменением: при определении продолженного пути вместо суммы длин начального участка и добавляемого ребра записывается максимальное из этих же двух чисел.

3) Определяется максимальная частота рёбер F_p в найденном на шаге 2 пути.

4) Если $F_p = F_{\max}$, то переходим к шагу 5. В противном случае осуществляется модификация частот: число 1 прибавляется к частотам всех рёбер пути, найденного на шаге 3. Текущее выполнение основного этапа завершено; переходим к очередному выполнению основного этапа.

5) При выполнении условия $F_p = F_{\max}$ рёбра с максимальной частотой образуют разрез графа. Если бы это было не так, то в силу минимаксного критерия алгоритм Дейкстры нашёл бы путь, не проходящий по рёбрам с максимальной частотой.

а) Находим компоненты связности исходного графа, которые получились бы при удалении всех рёбер с максимальной частотой.

б) Компоненту с максимальным числом вершин объявляем 1-ой частью текущей дихотомии графа, а другую компоненту (если она одна) или объединение всех остальных компонент (если их больше одной) объявляем 2-ой частью текущей дихотомии.

в) Запоминаем найденную дихотомию.

г) Как и в предыдущем случае, осуществляется модификация частот: число 1 прибавляется к частотам всех рёбер пути, найденного на шаге 3. Кроме этого, увеличивается и максимальная частота в рёбрах: $F_{\max} = F_{\max} + 1$. Текущее выполнение основного этапа завершено; переходим к очередному выполнению основного этапа.

Некоторые пояснения и иллюстрации к предложенному алгоритму приведены в статье [299].

Вернёмся к общей структуре алгоритма. Перед его выполнением задаётся число M требуемых дихотомий. Алгоритм останавливается, когда построено заданное число дихотомий.

5.2.3 Сцепленность графа

Для каждого значения единственного параметра M, результатом работы описанного в разделе 1 алгоритма будет семейство M дихотомий, т.е. разбиений исходного графа на две части. Для многих графов все найденные дихотомии могут оказаться совпадающими, но это неверно в общем случае. Напомним, что в алгоритме многократно используется датчик случайных чисел для последовательного выбора пар вершин. Поэтому в некоторых случаях, особенно практически важных, различные прогоны алгоритма производят к не всегда совпадающим дихотомиям. Эти случаи не являются досадными ошибками или недостатками метода. Более того, можно утверждать, что различные дихотомии естественно возникают при исследовании многих сложных систем, особенно тех, чьё функционирование определяется человеческой деятельностью. Следовательно, представляется целесообразным рассмотреть – как важное обобщение обычной задачи дихотомии – построение семейства дихотомий вместо единственной дихотомии. Построенное семейство характеризует исходную систему, моделируемую графом. Более того, в таких ситуациях сами

по себе отдельные дихотомии, формирующие указанное семейство, не вызывают интереса. Оказалось более целесообразным сосредоточить внимание на свойствах построенного семейства дихотомий, как целостного объекта.

Введём необходимые понятия и определения, позволяющие описать интересующие нас свойства построенного семейства дихотомий. Семейство всех дихотомий, найденных ЧАД, состоит из M элементов. Некоторые из них могут совпадать, а некоторые – быть различными. Пусть среди M дихотомий встречается всего D различных. Понятно, что в более простых случаях $D = 1$. В то же время в сложных случаях почти все построенные дихотомии могут оказаться различными.

Интуитивно понятно, что число D различных дихотомий, точнее, отношение этого числа к общему числу построенных дихотомий M, отражает некоторые свойства исходных графов и, следовательно, свойства рассматриваемых систем, моделируемых этими графами. Упомянутые свойства могут иметь различные содержательные интерпретации, но, в общем, они описывают сложность, запутанность, неясность, неопределенность и другие трудно формализуемые, хотя и весьма важные свойства разнообразных систем.

Такого рода величины в цитированных выше публикациях назывались индексами сложности (графа), декомпозиционной и дихотомической сложностью (графа), сцепленностью (графа). Просто для определённости в рамках данного отчёта будем пользоваться термином «сцепленность графа».

В ЧАД используется случайный датчик, поэтому результаты вычисления сцепленности для одного и того же графа могут быть различными, поскольку они являются реализациями некоторой случайной величины (определяемыми в результате различных прогонов ЧАД). Формально под сцепленностью графа понимается математическое ожидание этой случайной величины. В данном случае такое определение корректно просто в силу конечности множества всех возможных значений данной случайной величины (от 1 до M). Незначительность отличий различных реализаций сцепленности будет продемонстрирована в примерах, относящихся к фондовому рынку S&P-500.

Более подробно понятие сцепленности описано в отчёте МЛАВР 2018 года, а также в публикациях [296, 299].

5.2.4 Вычисление сцепленности для фондового рынка

Хотя все введённые выше понятия носят общий характер и относятся к произвольным неориентированным графам, специальные – так называемые рыночные графы, порождённые фондовыми рынками, в этом разделе рассмотрены подробнее.

Рассмотрен рынок S&P-500 (500 крупнейших компаний в США). Прежде всего дадим описание хорошо известной графовой модели произвольного фондового рынка. Объекты соответствуют рассматриваемым (в течение некоторого периода) акциям. Расстояние между двумя акциями определяется следующим образом.

1) Определим базисный минимальный период, состоящий из l последовательных дней. Все данные, найденные для периода x, $x-1$, **…** , $x-l+1$, сопоставляются дню x. Положим длину l рассматриваемого периода равной 16. Такой выбор определен содержательными соображениями: для короткого периода данные слишком вариативны, для длинного периода – слишком сглажены.

2) Цены всех акций на момент закрытия биржи рассмотрены для дней x, $x-1$, **…** , $x-l+1$. Матрица R попарных коэффициентов корреляции считается, исходя из этих цен.

3) Расстояния d_{ij} между двумя акциями (например, i и j) определяется формулой $d_{ij} = 1-r_{ij}$, где r_{ij} − соответствующий элемент матрицы R. Указанное расстояние d близко к 0 для «очень похожих» акций и близко к 2 для «очень непохожих» акций. Поэтому матрица $D = (d_{ij})$ рассматривается как матрица непохожести.

4) Устанавливается взаимно-однозначное соответствие между вершинами графа и рассматриваемыми акциями. Для каждого объекта a все остальные вершины упорядочиваются в порядке возрастания (неубывания) расстояния до a. Определяются первые четыре вершины в этом списке, а также все

вершины, расстояния от которых до *a* совпадают с расстоянием от четвертого в этом списке объекта. Вершина, соответствующая объекту *a*, соединяется ребрами со всеми этими вершинами. Легко видеть, что при указанном построении граф не зависит от конкретной нумерации объектов в упорядоченном указанным образом списке. Число ближайших вершин (здесь 4) является параметром при построении графа.

На первый взгляд, описанная хорошо известная конструкция не вызывает никаких вопросов. Предположим теперь, что начальный день *x* – это сегодняшнее число. Это означает, что ничего о поведении биржи в последующем неизвестно. Заметим, что это относится не только к курсам акций, но и к самому составу S&P-500. Зато о прошлом, естественно, известно всё.

Поэтому для построения рыночного графа для текущего (сегодняшнего) дня *x* можно взять все акции из сегодняшнего списка S&P-500, которые существуют на этом рынке не менее 16 дней. Поскольку изменение списка S&P-500 случается достаточно редко (5 – 15 раз в год), то таких акций обычно бывает столько же, сколько в день *x*. В редких случаях их бывает на одну или две меньше.

Описанный выбор акций хорошо подходит для постоянного наблюдения за курсами. Однако для рассматриваемого в отчёте большого периода в 30 лет такой анализ прошлого потребует огромного объёма данных, поскольку каждый день войдёт в 16 списков (напомним, что длина базисного периода $l = 16$) и соответствующего огромного объёма скачивания платных данных из финансовой базы данных Блумберга. Объём вычислений при этом также становится значительным. Вычисления для одного дня требуют 2-3 минут (при разумном числе испытаний $M = 2000$), не считая времени на подготовку данных.

Поэтому для однократного анализа данных за длительный период был предложен другой подход. В качестве начальных дней были рассмотрены последние дни каждого квартала: 31 марта, 30 июня, 30 сентября и 31 декабря. Скачивались данные, начиная с одного из этих дней, на квартал + ещё 15 дней назад. То есть данные с 31-го декабря по 16-ое сентября, с 30 сентября по 16-

ое июня, с 30-го июня по 17-ое марта, с 31-го марта по 17-ое декабря и т. д. Рассматривались все акции, торговавшиеся в последний день квартала и присутствовавшие ранее на рынке непрерывно начиная с указанных выше дней, т. е. в течение 107, 107, 106 и 105 дней подряд. Для тех дней, когда биржа не работает, все данные берутся из последнего рабочего дня.

Обычно такие наборы состояли из 496 – 500 акций. После 2010 года рассматривалось уже не 500, а 505 акций; соответственно, в указанные периоды на рынке было 501 – 505 акций.

Рассмотрим имеющийся набор N акций. Поскольку в течение квартала + 15 дней все они присутствовали на рынке, для любого дня квартала (92, 91 или 90 дней) известны их курсы на момент закрытия в этот день и ещё в 15-и предшествующих дней. Именно для этого и рассматривались 15-и дневные хвосты. Для любого дня x квартала и любых двух акций a и b из указанного списка можно подсчитать коэффициент корреляции по данным за 16 последовательных дней. Далее по всем попарным коэффициентам корреляции можно указанным в начале этого раздела способом определить рыночный граф и найти его сцепленность $D(x)$. Таким образом, сцепленность определяется для каждого дня x рассматриваемого квартала.

Центральным вопросом здесь является следующий. Предположим, мы нашли $D(x)$ для дня x, исходя из указанных выше 496 акций. Предположим, мы считаем теперь $D(x)$ для того же самого дня x, исходя из тех 500 акций, которые были на рынке в день x и в течение предшествующих 15-и дней. Конечно, все 496 акций участвовали в этих вычислениях, но, кроме, них, есть ведь и другие 4 акции. Не могут ли они заметно повлиять на сцепленности графов, определённых указанными выше способами и отличающихся 3-4-мя дополнительными вершинами?

К счастью, ответ на этот вопрос отрицательный. Изменения есть, но они не больше, чем изменения при разных прогонах алгоритма на одном и том же рыночном графе. Конечно, полная проверка для всех дней x не делалась, но во

всех случаях, считая и дни, непосредственно предшествующие большому кризису 2008 года, значения изменялись мало, и большие значения оставались большими, а малые значения – малыми. Подробнее эти данные будут приведены в следующем пункте.

Данный раздел завершается уже упомянутыми результатами вычисления сцепленности по отдельным кварталам за период 01.01.1991 – 30.09.2019. Они сведены в Таблицу А.1 Приложения А. Сцепленности приведены для каждого дня этого периода.

5.2.5 Анализ сцепленности для фондового рынка

На первый взгляд, данные в Таблице А.1 не обладают никаким очевидным порядком. Возвращаясь к определению сцепленности графа, можно сказать, что большие значения сцепленности означают хаотический характер графа, отсутствие каких бы то ни было кластеров вершин и регулярных связей между ними и пр. В то же время сравнительно малые значения сцепленности графа связаны с какими-то структурами на множестве его вершин, повторяемостью – и, следовательно, выделяемостью, хотя и не очень чёткой, – некоторых кластеров, словом, меньшей хаотичностью. Ещё раз стоит подчеркнуть, что понятие сцепленности связано с трудно определяемым, но достаточно важным понятием сцепленности, запутанности, сложности графа и, в силу этого, гораздо более важным аналогичным понятиям, относящимся уже не к графу, а к моделируемой им системе.

Применительно к определённым в начале предыдущего пункта рыночным графам можно сказать, что они в какой-то степени описывают динамику фондового рынка в течение 16 дней. Большое значение сцепленности, особенно в течение нескольких дней подряд, указывает на разнонаправленное и достаточно хаотичное изменение цен, хотя и не обязательно связанное с кризисными явлениями. А сравнительно небольшие значения сцепленности, также в течение нескольких дней подряд, предполагают бóльшую стабильность функционирования фондового рынке в эти периоды.

Рассмотрим теперь некоторые формальные особенности данных из Таблицы А.1. Прежде всего, будем считать значения сцепленности, превосходящие 1000, большими, а значения сцепленности, меньшие 100 – маленькими. Напомним, что все значения сцепленности изменяются в пределах 1 – 2000. Легко видеть, что больших значений в несколько раз меньше, чем маленьких. Ещё более различается их поведение в смысле образования серий. Для больших значений имеется всего 4 серии длины более 5. Они таковы:

01.01.1992 – 06.01.1992 со значениями 1797 1801 1823 1691 1789 1573

31.07.2008 – 07.08.2008 со значениями 1060 1250 1722 1220 1703 1250 1322 1254

05.08.2011 – 10.08.2011 со значениями 1420 1592 1227 1152 1785 1818

16.10.2011 – 22.10.2011 со значениями 1008 1795 1687 1824 1868 1423 1016

Для маленьких значений сцепленности ситуация совсем другая. Имеется 12 серий длины 15 и более. Они таковы:

07.03.2006 – 26.03.2006 со значениями 36 2 5 5 5 4 5 3 2 2 4 9 14 2 11 8 5 57 71 47

11.07.2008 – 25.07.2008 со значениями 30 17 16 13 10 5 7 12 6 7 5 9 4 7 15

23.11.2012 – 13.12.2012 со значениями 34 42 52 10 5 18 5 3 3 56 22 93 77 58 25 30 19 12 70 32

24.01.2014 – 07.02.2014 со значениями 48 45 9 13 11 16 17 70 29 11 10 11 32 23 48

11.11.2014 – 26.11.2014 со значениями 2 13 4 4 5 5 4 3 4 3 15 3 3 3 3 5

09.12.2014 – 24.12.2014 со значениями 5 2 3 2 3 5 2 7 5 4 2 2 2 2 2 4

10.02.2015 – 25.02.2015 со значениями 5 22 11 15 16 19 21 19 21 87 16 95 20 32 12 73

12.05.2015 – 27.05.2015 со значениями 3 2 3 2 2 3 3 3 3 3 16 3 2 3 3 9

09.06.2015 – 24.06.2015 со значениями 5 2 6 2 4 5 3 4 3 5 2 2 2 5 5 6

07.01.2016 – 21.01.2016 со значениями 11 37 17 14 4 9 5 5 4 7 31 39 5 19 31

07.01.2018 – 21.01.2018 со значениями 92 10 19 37 63 7 24 19 35 9 25 56 13 39 7

01.09.2019 – 18.09.2018 со значениями 23 10 4 7 10 31 23 43 5 28 6 9 27 34 56 29 31 44

Можно заметить, что в пределах одного квартала участков с большой и малой сцепленностью нет, кроме одного случая: периода 31.07.2008 – 07.08.2008 (серия из больших значений сцепленности) и почти сразу вслед за

ним, с 11.07.2008 по 25.07.2008 серия из малых значений сцепленности. Обратим внимание, что именно в этом же квартале примерно через месяц началась острая фаза ипотечного кризиса – самого крупного кризиса за последние тридцать лет.

5.2.6 Некоторые выводы

Приведенные выше соображения позволяют сформулировать некоторые направления дальнейших исследований, в конечном счете, направленных к пониманию функционирования фондового рынка. Конечно, для этого потребуется привлекать материалы, связанные не только с математическими моделями рынка, но и с содержательным описанием возникающих на нём эффектов и проблем.

Другим направлением работы предполагается исследование графов голосований различных политических органов. Как и для фондового рынка, получение необходимой информации о голосованиях является сложной неформальной задачей, требующей специальных усилий.

6 Исследование моделей распространения и специализации

В этом разделе отчета представлены результаты по исследованию моделей распространения и специализации. Исследования проходили в двух направлениях:

– исследование моделей распространения;

– исследование математических моделей возникновения специализации в многоклеточных системах.

6.1 Исследование моделей распространения

6.1.1 Введение

В настоящее время основным подходом к моделированию процесса распространения веществ в пространстве являются динамические математические модели, основанные на системах дифференциальных уравнений. Тем не менее, ни одна из моделей не в состоянии ответить абсолютно всем потенциально возможным ситуациям по той причине, что всегда остаётся определенное количество некоторых физических параметров, которые влияют на процесс, однако остаются неучтёнными.

В свою очередь, теория устойчивости разностных схем для модельных задач газовой динамики хорошо представлена в литературе. Впервые рассмотрены конечно-разностные схемы решения одномерных уравнений газовой динамики для случая полной системы из законов сохранения массы, импульса и энергии. Эти схемы являются явными двухуровневыми по времени и используют симметричный трёхточечный шаблон по пространственной оси. Их построение основано на специальной квазигазодинамической и квазигидродинамической [300, 301] регуляризации исходного уравнения (без регуляризации схемы нестабильны). Схемы такого рода успешно применяются в многочисленных практических приложениях, однако их теория до сих пор была недо-

статочно развита. В одномерном случае впервые в [302, 303, 304] были выведены критерий и как необходимое, так и достаточное условие линеаризованной L^2-диссипативности, а также проведён анализ применимости данных разностных схем в применении к нелинейной постановке для произвольной фоновой скорости.

Помимо этого, было решено уделить внимание прикладным задачам, напрямую связанным с распространением веществ. В частности, была рассмотрена задача оптимального расположения сенсоров, которое позволит в максимально короткие сроки обнаруживать возникающее нефтяное пятно на поверхности моря и оперативно реагировать на данную угрозу.

Безусловно, в идеальной ситуации надо поставить датчики на поверхности моря максимально плотно. Однако, поскольку всё это требует денежных затрат, а бюджет на превентивные меры против чрезвычайных ситуаций ограничен, возникает ситуация, при которой надо максимизировать площадь, покрываемую сенсорами, при этом минимизируя экономические издержки на установку и поддержку работы датчиков.

Для решения данной проблемы чаще всего используется линейное программирование [305, 306]. Однако, в классической постановке задачи все участки, за которыми надо следить, равнозначны. Однако, в случае с возникновением разлива нефти на поверхности моря ситуация отличается.

Некоторые факторы, такие как морские течения и ветра, позволяют достаточно точно узнавать значение тех или иных параметров при небольшом смещении в координатах в том же направлении, что и вышеупомянутые факторы.

В результате, можно утверждать, что существуют некоторые участки моря, значения параметров в которых позволят с большой точностью вычислить значение этих же параметров в других участках, даже находящихся за пределами действия сенсора. Для обозначения таких участков предлагается использовать термины «критические» или «ключевые» точки.

Благодаря этому, в рамках проекта было предложено модифицировать задачу. Ведь теперь не все участки одинаково важны для покрытия сенсорами. Следовательно, проблема разбивается на два основных этапа – выделение ключевых точек и оптимизация покрытия сенсоров с учётом расположения критических точек.

6.1.2 Исследование решений одномерных уравнений газовой динамики

В рамках исследования динамических моделей процессов распространения на основе анализа систем дифференциальных уравнений была рассмотрена задача одномерной газовой динамики

$$\partial_t \rho + \operatorname{div} \rho u = 0, \partial_t(\rho u) + \operatorname{div}(\rho u^2) + \nabla p = \operatorname{div} \Pi_{NS},$$

$$\partial_t E + \operatorname{div}[(E + p)u] = \partial_x(\varkappa \partial_x \varepsilon) + \partial_x(\Pi_{NS} * u),$$

где искомые плотность $\rho > 0$, скорость газа u и внутренняя энергия $\varepsilon > 0$ зависят от *(x,t)*, $\varkappa$ – масштабированный коэффициент теплопроводности.

В результате исследований были линеаризованы регуляризованные явные конечно-разностные схемы на постоянном решении для данной задачи относительно плотности, скорости газа и внутренней энергии. При этом полная энергия E задаётся формулой. $E = 0.5\rho u^2 + \rho\varepsilon$, а давление – уравнением состояния $p = p(\rho, \varepsilon)$. Как и в баротропном случае, в теплопроводном случае было получено необходимое условие для слабой консервативности $\left(\sup_{m \geq 0} \|y^m\|_H \leq \|y^0\|_H \ \forall y^0 \in H\right)$.

В общем случае для абстрактной явной разностной схемы вида $\delta_t \vec{y} - c_* B \mathring{\delta} \vec{y} + \tau_* c_*^2 A \delta \vec{y} = 0$ с некоторыми комплекснозначными матрицами A и B были выведены критерий, а также как необходимое, так и достаточное условие типа CFL для слабой консервативности в L^2 для соответствующей начальной задачи. Слабая консервативность в L^2 означает равномерность временной границы для нормы масштабированного решения вместо закона сохранения энергии для линеаризованной исходной системы уравнений. Впервые абстрактный

подход был исследован в [304] для баротропной системы, где предполагалось, что матрицы A и B порождают самосопряжённые линейные операторы, однако затем в теплопроводном случае [303] соответствующие условия были выведены и в общем случае, где матрицы A и B, вообще говоря, не являются эрмитовыми. Тем не менее, в линеаризованной полной системе соответствующие матрицы, отвечающие за вязкие и конвективные слагаемые, являются самосопряжёнными, что позволяет получить условия в виде, удобном для практического использования.

Кроме того, в [303] был проведён ряд численных экспериментов на модельной задаче Римана, в ходе которых было апробировано три различные дискретизации, эквивалентные с точки зрения линеаризации, однако различные по своему характеру. В частности, две схемы, предложенные ранее проф. Злотником А.А., обладают свойством так называемой энтропийной диссипативности. Тщательные эксперименты показали, что энтропийно-диссипативные схемы не только позволяют получать более широкие диапазоны числа Куранта, что ускоряет время вычислений, но также показывают достаточно хорошее соответствие с необходимым условием устойчивости линеаризованной схемы. Тем самым, показано, что граница применимости условий, полученных путём линеаризованного анализа, не ограничивается только лишь линейным случаем, и для обладающих достаточно хорошими свойствами разностных методов можно применять данные условия и в нелинейной постановке для реальной задачи.

6.1.3 Задача оптимального расположения сенсоров обнаружения загрязнений на поверхности моря

Что касается оптимизации расположения сенсоров на поверхности моря, для реализации первого этапа предложенной модели весь исследуемый участок моря разбивается на элементарные территории небольшой площади. Из этих элементарных территорий путём соединения граничащих друг с другом участков формируется сеть.

Затем на основе данной сети строится целая группа однородных сетей, в каждой из которых веса на рёбрах выставляются на основе одного из заранее выбранных факторов, который активно влияет на перемещение между элементарными территориями. Так, например, строится отдельная сеть перемещения за счёт морских течений. Для этого, используя усреднённую по времени карту морских течений, для каждой пары элементарных территорий i и j вес ребра, идущего из i в j, устанавливается равным проекции вектора средней скорости течения на данное направление. Затем, аналогичным образом строится сеть и для всех остальных выбранных факторов (ветра, морские пути, …).

После этого, для анализа всех построенных сетей используются основные методы сетевого анализа, позволяющие оценить влияние отдельных вершин в сети. В первую очередь, речь идёт о таких классических индексах центральности, как eigenvector [307] и PageRank [308].

Eigenvector центральность назначает относительные оценки важности всем узлам в сети. В основе этого подхода лежит идея, согласно которой соединения с узлами с высокой оценкой важности вносят больший вклад в оценку рассматриваемого узла, чем аналогичные соединения с узлами с низкой оценкой важности.

Для практического подсчёта данного индекса центральности составляется матрица смежности сети A. Затем, решив линейную систему $Av=v$, значение eigenvector центральности для i-ой вершины в сети будет равно i-ой координате вектора v.

В свою очередь для вычисления индекса центральности PageRank производится симуляция случайного марковского процесса – блуждания по сети, при котором вероятность перехода из состояния i в состояние j равна весу ребра, исходящего из i в j, нормированная на сумму весов всех исходящих из вершины i рёбер. Кроме того, в этом блуждании заложена некоторая вероятность окончания блуждания с последующим повторным запуском из случайной вершины сети. В результате данного моделирования вычисляется среднее

время нахождения в каждой вершине, которое и даёт значение индекса центральности PageRank.

Когда все индексы на всех сетях посчитаны, остаётся выделить критические точки. Для начала надо выделить наиболее важные вершины сети. Для этого было решено использовать фиксированное количество лучших вершин с точки зрения каждого индекса для каждой сети (лучшие 10 вершин). Альтернативным решением может быть использование всех вершин, индекс центральности которых превышает заданный заранее порог (например, все вершины со значением индекса центральности выше 0.5).

После выделения нескольких наборов вершин, признанных важными с точки зрения одного из индексов центральности на одной из рассматриваемых сетей, остаётся лишь агрегировать эти наборы. Для этого также есть много разных подходов, однако было выбрано самое простое решение – просто объединить все эти множества. Таким образом, ключевыми становятся все вершины, попавшие хотя бы один раз в список 10 лучших согласно одному из индексов центральности на одной из рассматриваемых сетей.

Однако, если использовать только вершины, имеющие высокое значение индекса центральности, есть вероятность, что все выделенные вершины будут сконцентрированы в одном месте и большая часть моря не будет покрыта. Например, такая ситуация возникает в случае анализа некоторого пролива, имеющего «бутылочное горлышко» в середине. В этом случае все критические точки будут находиться рядом с этим «бутылочным горлышком» поскольку именно через них будет проходить большинство кратчайших путей.

Для решения этой проблемы предлагается добавить в множество критических точек нефтедобывающие вышки и/или пути прохождения танкеров. Это поможет следить за всеми локациями, представляющими опасность, - как стационарными, так и передвижными.

На этапе оптимизации покрытия критических точек, как уже упоминалось при описании общей структуры модели, решается задача выбора некоторого подмножества всех критических вершин, которое удовлетворяет всем

наложенным на модель ограничениям. В результате решается следующая задача:

$$\begin{cases} \sum_{i=1}^{N} x_i \rightarrow max, \\ \sum_{i=1}^{N} x_i\, c_i \leq S, \end{cases}$$

где N – общее количество критических точек, x_i – бинарная переменная, показывающая надо ли ставить в i-й критической точке сенсор, c_i – суммарная стоимость обслуживания сенсора в i-й критической точке (установка и поддержание стабильной бесперебойной работы), а S – бюджет, выделенный на решение данной задачи.

6.1.4 Некоторые выводы

По итогам работы можно сказать, что поставленные задачи выполнены в полном объёме. Во-первых, что теоретически были исследованы явные двухуровневые конечно-разностные схемы решения полной системы уравнений газовой динамики для случая произвольной фоновой скорости.

Выведенные условия являются достаточно широкими, и при этом, несмотря на свою линеаризованную природу, численные эксперименты показывают хорошее соответствие с прикладными расчётами в исходной нелинейной постановке.

Во-вторых, была получена универсальная модель, которая может применяться к территории любого масштаба (участку моря, целому морю, всему Северному Ледовитому океану) для нахождения оптимального расположения сенсоров в море [309]. Данная модель анализирует рисунок морских течений, ветров и прочих факторов, влияющих на распространение веществ на поверхности моря. Как следствие, упрощается задача по оптимальному расположению сенсоров благодаря меньшему перебору возможных локаций для них.

При этом данная модель является достаточно простой в применении и при необходимости может подстраиваться и усложняться за счёт добавления дополнительных ограничений на этапе оптимизационной задачи или же при изменении правил выбора критических точек.

6.2 Исследование математических моделей возникновения специализации в многоклеточных системах

6.2.1 Введение

Системы из конечного числа элементов, каждый из которых может выполнять одну или несколько функций так, что выполнение одной из них ведет к снижению выполнения хотя бы одной другой, встречаются в различных областях знаний – в различного рода биологических системах, на различных уровнях организации живой материи, в экономике, социальных науках. Системы описанного типа будем называть многоклеточными. Проблема возникновения функциональной специализации в многоклеточных системах фундаментальна. Актуален вопрос выявления факторов, влияющих на возникновение специализации в многоклеточных системах, и развития аппарата для ее исследования. Так, в эволюционной биологии возникновение дифференцированных типов клеток, то есть, многоклеточной специализации, ведет к структурному усложнению организмов и увеличению биологической сложности. В эволюционной антропологии зарождение в первобытных сообществах разделения труда - это двигатель к развитию экономики.

В социально-экономических науках понятие специализации исторически носит название «разделение труда». Это понятие изучал в своих фундаментальных трудах Адам Смит [310], полагая, что разделение труда представляет собой ключевое условие для увеличения производительности труда, прогресса производственных сил, развития экономики, общества и государства. Современные работы в области экономической теории продолжают развитие этих идей, исследуя влияние разделения труда на экономическое развитие,

рост и прогресс [311]. Огромный пласт работ посвящен возникновению специализации в социальных науках, а точнее – вопросу возникновения разнообразия или неоднородности в системах. В математической теории социальных институтов и организаций исследуется разделение труда между элементами различных социальных институтов и его связь с возникновением внутри этих систем иерархической структуры [312]. В социальной антропологии изучается, каким образом в первобытных сообществах могли формироваться первые иерархические структуры и социальные институты [313]. Предлагаемые математические модели описывают также природу альтруизма, кооперации и механизмов подавления конфликта в первобытных сообществах [313, 314].

Развитие жизни на земле можно описать в виде последовательности фундаментальных переходов от набора объектов низкого уровня к одному объекту более высокого уровня организации живой материи [315]. Фундаментальны также переход от автокаталитических реакций к РНК-миру [316], от генов к прокариотической клетке, от прокариотической клетки к эукариотической клетке, от одной клетки к многоклеточному организму, от единичного организма к сложным социальным структурам. Эти переходы осуществляются благодаря возможности объектов более низкого уровня специализироваться на выполнении различных функций, присущих объекту более высокого уровня; возможности к кооперации и альтруизму; подавлению конфликта внутри объекта более высокого уровня.

Поэтому возникает необходимость в разработке математических моделей возникновения специализации в многоклеточных системах. Приведем далее описание модели и основных ее результатов.

Рассмотрим абстрактную систему из N элементов. Пусть $i = 1, \dots, N$ – индексы элементов в системе. Предположим, что каждый элемент может выполнять одну из двух задач. Обозначим через $b_i \in [0, b_i^{max}]$ уровень выполнения первой задачи элементом с индексом i, через $v_i \in [0, v_i^{max}]$ – уровень выполнения второй задачи элементом с индексом i. Введем обозначения $b = (b_1, \dots, b_N)$ и $v = (v_1, \dots, v_N)$.

Теперь определим уровни выполнения задач всей системой целиком. Мы будем полагать, что уровень выполнения системой первой задачи (B) есть аддитивная функция переменных b_i, а уровень выполнения системой второй задачи (V) есть аддитивная функция переменных v_i:

$$B(b) = \sum_{i=1}^{N} b_i\,; \quad V(v) = \sum_{i=1}^{N} v_i\,. \tag{6.1}$$

Определим индекс эффективности системы (W) как произведение уровней выполнения системой первой и второй задач:

$$W = V \cdot B. \tag{6.2}$$

Мы используем данный индекс, следуя работе [317], и исходя из того, что данный индекс отражает синергетический эффект, возникающий между элементами одной системы.

Будем предполагать, что внутренняя структура каждого элемента такова, что элемент не может увеличивать выполнение одной из двух задач, не уменьшив выполнения другой. Данная предпосылка формализуется понятием структурного ограничения элемента:

$$v_i = v_i^{max} - \alpha_i \cdot b_i, \tag{6.3}$$

где α_i – неотрицательные коэффициенты.

Будем говорить, что два элемента системы имеют одинаковые типы, если соответствующие параметры их структурных ограничений совпадают [318, 319, 320].

Для каждого элемента с индексом $i \in \{1, \dots, N\}$, обозначим через b_i^{max} величину $b_i^{max} = \frac{v_i^{max}}{\alpha_i}$. Линейность структурных ограничений позволяет упростить дальнейшие рассуждения.

Мы будем предполагать также, что внешняя среда оказывает влияние на эффективность функционирования системы. Пусть C – это уровень ресурса, доступный системе при заданных параметрах внешней среды. Предположим, что для выполнения единицы первой задачи системе требуется k_1 единиц ресурса, а для выполнения единицы второй задачи – k_2 единиц ресурса. Кроме того, если некоторый элемент системы выполняет одновременно обе задачи, это влечет за собой наличие дополнительных затрат ресурсов (связанных, например, с необходимостью пространственного разделения процессов выполнения различных задач внутри элемента). Таким образом, общее ресурсное ограничение имеет вид:

$$k_1 \sum_{i=1}^{N} b_i + k_2 \sum_{i=1}^{N} v_i + k_3 \sum_{i=1}^{N} b_i \cdot v_i \leq C, \text{где } k_1, k_2, k_3, C > 0. \quad (6.4)$$

Рассмотрим выражение $\sum_{i=1}^{N} b_i \cdot v_i$. В данной сумме слагаемое с индексом i может быть рассмотрено как индекс эффективности системы, состоящей из одного единственного элемента с индексом i, вычисленный согласно формуле (6.2). Иначе говоря, слагаемое $\sum_{i=1}^{N} b_i \cdot v_i$ представляет собой суммарную эффективность заданного набора элементов вне рассматриваемой системы. Поэтому будем называть слагаемое $k_3 \sum_{i=1}^{N} b_i \cdot v_i$ издержками оппортунистического поведения элементов системы. Можно предположить, что ненулевые значения параметра k_3 могут возникнуть в процессе эволюции системы с целью поддержания стабильности этой системы.

Оптимальной стратегией системы мы будем называть набор (b^*, v^*) таких уровней выполнения задач каждым из элементов, которые при условии существования структурных ограничений и общего ресурсного ограничения

позволяют максимизировать эффективность функционирования системы. Таким образом, мы можем свести нашу задачу к классической задаче математического программирования с ограничениями:

$$\begin{cases} W = \sum_{i=1}^{N} b_i \cdot \sum_{i=1}^{N} v_i \to max_{b,v} \\ \forall i = 1..N: v_i = v_i^{max} - \alpha_i \cdot b_i, \\ k_1 \sum_{i=1}^{N} b_i + k_2 \sum_{i=1}^{N} v_i + k_3 \sum_{i=1}^{N} b_i \cdot v_i \le C, \\ \forall i = \overline{1,N}: v_i \ge 0, \\ \forall i = \overline{1,N}: b_i \ge 0. \end{cases} \tag{6.5}$$

Или иначе:

$$\begin{cases} W = \sum_{i=1}^{N} b_i \cdot \left(\sum_{i=1}^{N} v_i^{max} - \sum_{i=1}^{N} \alpha_i \cdot b_i\right) \to max_b \\ k_1 \sum_{i=1}^{N} b_i + \sum_{i=1}^{N} (k_2 + k_3 b_i) \cdot (v_i^{max} - \alpha_i \cdot b_i) \le C \\ \forall i = \overline{1,N}: 0 \le b_i \le b_i^{max}. \end{cases} \tag{6.6}$$

Будем говорить, что элемент с индексом i специализируется на первой задаче, если $b_i^* = b_i^{max}$(равносильно $v_i^* = 0$). Аналогично, будем считать, что элемент с индексом i специализируется на второй задаче, если $v_i^* = v_i^{max}$(равносильно $b_i^* = 0$).

6.2.2 Основные результаты

Введем обозначение: $H = \{b \in R^N | \forall i = \overline{1,N}: 0 \le b_i \le b_i^{max}\}$. Рассмотрим следующее множество: $A(\overline{W}) = \{b \in H | \, W(b) \ge \overline{W}\}$.

Утверждение 6.1. При любом значении $\overline{W}$ множество $A(\overline{W})$ является выпуклым.

Утверждение 6.2.

1) ресурсное ограничение задачи (6.6) задает замкнутое множество в евклидовом пространстве, границей которого является (N-1)-мерный эллипсоид:

$$\sum_{i=1}^{N} \alpha_i \cdot (b_i - \gamma_i)^2 \geq \sum_{i=1}^{N} \alpha_i (\gamma_i)^2 + \frac{k_2 \cdot \sum_{i=1}^{N} v_i^{max} - C}{k_3},$$

где $\gamma_i = \frac{k_1 - k_2 \alpha_i + k_3 v_i^{max}}{2 k_3 \alpha_i}, i = \overline{1, N}$;

2) обозначим через D множество допустимых решений задачи (6.6). Тогда существует такое значение $C_1 > 0$, что при любом значении параметра C из множества $(0, C_1)$, множество D является пустым, а для всех $C \geq C_1$ множество D не пусто. Более того, данное значение C_1 может быть найдено следующим образом:

$$C_1 = k_1 \cdot \sum_{i \in S} b_i^{max} + k_2 \cdot \sum_{i \in \{1,..,N\} \setminus S} v_i^{max},$$

где $S = \left\{ i \in \overline{1, N} \middle| \alpha_i \geq \frac{k_1}{k_2} \right\}$.

Определим следующее множество:

$$G = \{ b \in H | k_1 B(b) + k_2 V(b) + k_3 \langle b, v(b) \rangle = C \}.$$

Обозначим через B^* множество оптимальных решений задачи (6.6). Обозначим через B^{**} множество оптимальных решений задачи (6.6) без ресурсного ограничения.

Утверждение 6.3. Существуют такие значения параметра C, C_1 и C_2, $0 < C_1 \leq C_2$, что:

1) $\forall C \geq C_2 : B^* \subseteq B^{**}$,

2) $\forall C \in [C_1, C_2) : B^* \subseteq G$,

3) $\forall C \in [0, C_1) : B^* = \emptyset$,

где $C_1 = k_1 \cdot \sum_{i \in S} b_i^{max} + k_2 \cdot \sum_{i \in \{1,..,N\} \setminus S} v_i^{max}$, $C_2 = min_{b \in B^{**}} \{ k_1 B(b) + k_2 V(b) + k_3 \langle b, v(b) \rangle \}$.

Таким образом, мы можем сделать следующие выводы относительно того, как устроено решение задачи (6.6). Во-первых, существует пороговое

значение уровня доступного системе ресурса, такое, что если содержание ресурса во внешней среде опустится ниже данного порогового значения, система перестанет функционировать вообще. Во-вторых, существует второе пороговое значение для уровня доступного ресурса. Если содержание ресурса во внешней среде выше данного значения, то ресурсное ограничение никак не влияет на эффективность системы – реализуется оптимальная стратегия при наличии только структурных ограничений. При любом уровне ресурса, находящемся в промежутке между двумя пороговыми значениями, ресурсное ограничение оказывает влияние на эффективность системы. При этом системой используется все доступное количество ресурса, что выглядит весьма реалистично в рамках практических интерпретаций рассматриваемых моделей.

6.2.3 Случай системы с элементами одного типа

Наибольший интерес для нас будет представлять случай модели, когда все элементы в системе имеют одинаковый тип. В таком случае ни у одного элемента системы нет никакой заранее заданной структурной предрасположенности к выполнению одной из задач относительно других элементов системы. Итак, предположим, что все элементы системы имеют одинаковый тип.

Утверждение 6.4. Справедливы следующие положения:

1) $C_2 = \frac{N}{2} \cdot (k_1 b^{max} + k_2 v^{max}) + k_3 \left(N - 2 \cdot \left[\frac{N}{2}\right]\right) \cdot \frac{b^{max} v^{max}}{4}$;

2) для всех значений параметра C из отрезка $[C_1, C_2]$ в оптимуме будет не более одного неспециализированного элемента;

3) существует значение $C_3 = \frac{N}{2} \cdot (k_1 b^{max} + k_2 v^{max}) + k_3 N \cdot \frac{b^{max} v^{max}}{4}$, такое что для всех $C \in [C_2, C_3)$: $B^* \subset B^{**}$, а для всех $C \geq C_3$: $B^* = B^{**}$, где $B^{**} = \left\{b \in H \mid B(b) = \frac{N b^{max}}{2}\right\}$.

Таким образом, мы можем сделать следующие выводы из приведенного выше утверждения. Во-первых, если ресурсное ограничение влияет на оптимальную эффективность системы и позволяет системе существовать, то все

элементы системы кроме, возможно, одного будут специализированными. Во-вторых, при всех $C \geq C_3$ множество решений задачи (6.6) совпадает с множеством решений задачи (6.6) без ресурсного ограничения – связным множеством стратегий, где стратегии, в которых хотя бы один элемент является специализированным, образуют подмножество меры ноль. При уменьшении C от C_3 к C_2 происходит уменьшение множества оптимальных решений до тех пор, пока мы не придем к ситуации, когда в оптимуме будет не более одного неспециализированного элемента. Мы видим, что по мере уменьшения количества доступного ресурса, система будет вести себя следующим образом: сначала система будет оставаться неспециализированной. Затем каждый элемент также будет оставаться неспециализированным, но уровни выполнения некоторых функций данным элементом будут доминировать уровни выполнения других, пока, наконец, каждый элемент не станет специализированным. В конечном итоге, система вымрет от недостатка ресурса. Для иллюстрации полученных результатов рассмотрим следующий пример.

6.2.4 Некоторые выводы

Предложены общие математические модели проблемы разделения труда между элементами некоторой абстрактной системы. Введено понятие специализации в рассматриваемых абстрактных системах, сформулирована оптимизационная задача максимизации эффективности функционирования системы. Изучены возможности возникновения специализации в оптимуме в зависимости от наличия в модели элементов различных типов и от влияния факторов окружающей среды (при предположении о линейности всех рассматриваемых ограничений). Приведены общие результаты, описывающие структуру решения рассматриваемой задачи математического программирования. Более детально рассмотрен случай, когда все элементы в системе идентичны. Для данного случая описано, как и при каких условиях возможно возникновение специализации, и как различные параметры модели влияют на переход системы к

специализированным состояниям. Рассматриваемые в работе абстрактные модели специализации могут быть применены для описания широкого круга проблем не только в эволюционной биологии, но и в социальных науках, истории, экономике.

ЗАКЛЮЧЕНИЕ

При выполнении проекта получены следующие основные результаты, могут быть сделаны следующие выводы и рекомендации:

– в рамках исследования манипулируемости правил агрегирования предпочтений проведены вычислительные эксперименты для расчета вероятности индивидуального манипулирования при наличии контругроз для 11 правил коллективного выбора; показано, что правилами подсчета очков с одним скачком баллов всегда можно манипулировать; доказано, что для правила Борда небезопасно манипулируемый профиль существует тогда и только тогда, когда число избирателей или число альтернатив больше или равно 5;

– исследованы различные подходы и методы анализа решений в условиях неопределенности, основанные на теории важности критериев и не предполагающие использования функций полезности и количественных вероятностей;

– выполнен аксиоматический анализ способов выбора оптимальных вариантов и построения рейтинга альтернатив;

– в рамках исследования профилей предпочтений получены: точная формула для числа группоразделимых профилей предпочтений; рекурсивная формула для числа нарцисстических группоразделимых профилей предпочтений;

– в рамках исследования стабильных паросочетаний разработано новое расширение условия последовательного предпочтения, которое приводит к единственному стабильному паросочетанию для любого подмножества агентов; в новых условиях алгоритм Гейла-Шепли решения задачи о марьяже является стабильным, последовательным, стратегически устойчивым, оптимальным по Парето для мужчин и оптимальным по Парето для женщин;

– разработан метод оценивания влияния в мультиплексных сетевых структурах; предложенная модель применена к сетям международной торговли продовольствием и торговли вооружением;

– предложен и исследован ряд мер оценивания важности рёбер в сетевых структурах;

– проведен сетевой анализ международной миграции с 1970 по 2014 гг. на основе данных по потокам миграции между странами; определены рейтинги влиятельных стран в процессе международной миграции с помощью индексов центральностей, как классических, так и индексов ближних и дальних взаимодействий;

– проведен сетевой анализ международной торговли с учетом прямого и непрямого взаимодействия между странами;

– в рамках исследования методов анализа текстов вводится понятие наиболее подходящего обобщения нечеткого множества тем, сформированного листьями корневого дерева таксономии; предложенный метод обобщения “поднимает” нечеткое множество тем к узлам таксономии более высокого ранга наиболее “тесным” способом; метод апробирован на текстовой коллекции из 18000 научных публикаций по тематике анализа данных, и таксономии Науки данных, сформированной на основе Таксономии Computing Machinery Computing Classification System 2012, ACM-CCS;

– в рамках разработки методов анализа паттернов предложен алгоритм выделения симметричных паттернов на больших и сильно зашумленных данных; показано, что симметрия паттернов сохраняется при любом порядке их формирования;

– в рамках исследования конкурентной кооперации в Арктике и на основе результатов учебного и исследовательского курсов в рамках проекта «Общество и передовые технологии в Арктике» (SATA) введена и

описана концепция цикла SATA, которая отражает междисциплинарность применения фундаментальных исследований для улучшения условий жизни в Арктике с использованием передовых технологий;

– предложена математическая модель для оценки необходимого объема инвестиций в развитие региональной инфраструктуры грузоперевозок; предложенная модель апробирована на модельных примерах грузовых транспортных систем; формулируются некоторые рекомендации по использованию этой модели;

– выявлены и описаны классы системных проблем, связанных с формированием государственно-частного партнерства (на примере энергетических систем);

– модифицированы математические модели, на основе которых переформулированы математические задачи оценки объема инвестиций, требуемого для размещения цепи зарядных станций электромобилей на шоссе; доказаны соответствующие математические утверждения;

– показана возможность применения методов теории неаддитивных мер и функций доверия к решению ряда задач политологического анализа: нахождения наиболее значимых для голосования групп вопросов, оценивания политической неоднородности общества, оценивания значимости позиций отдельных партий для голосования, оценивания конфликтности позиций партий по значимым вопросам; исследование выполнено по базе данных интернет-опросов партий в Германии перед выборами в Бундестаг в 2013г. и результатов этих выборов;

– проанализированы основные модели и методы оценивания поляризации и политического влияния, в том числе и модели, учитывающие предпочтения по формированию коалиций;

– в рамках анализа моделей в банковской сфере проанализировано влияние модельного риска на точность оценок величины риск-взвешенных

активов, полученных с помощью подхода на основе внутренних рейтингов; предложены меры по учету такого влияния в задаче повышения сопоставимости оценок кредитного риска;

– в рамках исследования сруктурной сложности графов разработана новая модификация показателя сложности графа, основанная на построении семейства графовых дихотомий; с помощью предложенного подхода анализируются графы для данных фондового рынка S&P-500 за период 1991-2019гг. с целью прогнозирования особых периодов в поведении фондового рынка;

– исследованы явные двухуровневые конечно-разностные схемы решения полной системы уравнений газовой динамики для случая произвольной фоновой скорости;

– поставлена и исследована оптимизационная задача возникновения специализации в абстрактной многоклеточной системе на основе максимизации эффективности функционирования этой системы.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Отчет МНУЛ анализа и выбора решений, ТЗ-71 за 2018г.
2. Gibbard A. Manipulation of voting schemes: a general result// Econometrica: Journal of the Econometric Society. – 1973. – N 41(4). – P. 587 – 601.
3. Satterthwaite M.A. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions// Journal of economic theory. – 1975. – N 10(2). – P. 187 – 217.
4. Favardin P., Lepelley D. Some further results on the manipulability of social choice rules// Social Choice and Welfare. – 2006. – N 26. – P. 485 –-509.
5. Aleskerov F., Karabekyan D., Sanver R., Yakuba V. An individual manipulability of positional voting rules// SERIEs: Journal of the Spanish Economic Association. – 2011. – N 2. – P.431 – 446.
6. Aleskerov F., Kurbanov E. A Degree of Manipulability of Known Social Choice Procedures// Current Trends in Economics: Theory and Applications / Eds. Alkan A., Aliprantis Ch., Yannelis N. N.Y.: Springer-Verlag, 1999. P. 13 - 27.
7. Slinko A., White S. Is it ever safe to vote strategically?// Social Choice and Welfare. – 2014. – N 43(2). – P. 4 03 – 427.
8. Веселова Ю. А. Всегда ли безопасно коалиционное манипулирование результатом голосования?// В кн.: Управление развитием крупномасштабных систем (MLSD’2019): материалы Двенадцатой международной конференции, 1–3 окт. 2019 г. / Под общ. ред.: С. Н. Васильев, А. Цвиркун. Институт проблем управления им. В.А. Трапезникова РАН. – 2019. – С. 1185 – 1187.
9. Алескеров Ф.Т., Карабекян Д.С., Иванов А.А., Якуба В.И. О манипулируемости процедуры выбора Хара// В кн.: XIII Всероссийское совещание по проблемам управления ВСПУ-2019: труды. – М.: ИПУ РАН. – 2019. С. 1 – 6.
10. Карабекян Д.С., Иванов А.А., Якуба В.И. Манипулируемость схем агрегирования в модели Даунса// В кн.: Управление развитием крупномасштабных систем (MLSD’2019): материалы Двенадцатой международной конференции, 1–3 окт. 2019 г. / Под общ. ред.: С. Н. Васильев, А. Цвиркун.

Институт проблем управления им. В.А. Трапезникова РАН. - 2019. - С. 979 – 981.

11. Aleskerov F.T., Karabekyan D., Ivanov A., Yakuba V.I. On the bounds of weak manipulability of majoritarian aggregation procedures// International Conference on Information Technology and Quantitative Management (ITQM 2019), November, 3-6, 2019 (в печати).

12. Fishburn P.C. Decision and value theory. – New York: Wiley, 1964.

13. Подиновский В.В. Анализ решений при множественных оценках коэффициентов важности критериев и вероятностей значений неопределенных факторов в целевой функции// Автоматика и телемеханика. – 2004. – N 11. – С. 141 – 159.

14. Льюс Р.Д., Райфа Х. Игры и решения. Введение и критический обзор / Пер. с англ. – М.: Изд-во иностр. лит-ры, 1961.

15. Schoemaker P. The expected utility model: its variants, purposes, evidence and limitations// Journal of economic literature. – 1982. – N 20. – P. 529 – 563.

16. Ларичев О.И. Теория и методы принятия решений: Учебник. – М.: Логос, 2006.

17. Подиновский В.В. Аксиоматическое решение проблемы оценки важности критериев в многокритериальных задачах принятия решений// Современное состояние теории исследования операций / Под ред. Н.Н. Моисеева. – М.: Наука, 1979. – С. 117 – 145.

18. Подиновский В.В., Потапов М.А., Нелюбин А.П., Подиновская О.В. Теория важности критериев: современное состояние и направления дальнейшего её развития// XII Всероссийское совещание по проблемам управления ВСПУ-2014. Москва, 16 – 19 июня 2014 г. ИПУ РАН [Электронный ресурс]. – М.: ИПУ РАН, 2014. – С. 7697 – 7702. http://vspu2014.ipu.ru/node/8581

19. Подиновский В.В. Задачи принятия решений при качественном оценивании факторов// Методологические проблемы использования математических методов в социологических исследованиях. - М.: ИСИ АН СССР, 1980. – С. 115 – 125.

20. Подиновский В.В. Задачи принятия решений при качественном оценивании факторов и предпочтений// Математические методы в социологическом исследовании. - М.: Наука, 1981. – С. 115 – 125.

21. Podinovski V.V., Podinovski Vic.V. Decision analysis under partial information// Trends in multicriteria decision making / Th.J. Stewart, R.C. Van der Honert (Eds.). Lecture notes in economic and mathematical system. - Berlin: Springer, 1998. - N 465. - P. 77 – 84.

22. Пфанцагль И. Теория измерений. – М.: Мир, 1976.

23. Fishburn P.C. The axioms of subjective probability// Statistical Science. – 1986. – N 1. – P. 335 – 345.

24. Наумов Г.Е., Подиновский В.В. Подиновский Вик.В. Субъективная вероятность: способы представления и методы получения// Известия АН СССР. Техническая кибернетика. – 1991. – N 1. – С. 94 – 109.

25. Kraft C.H., Pratt J.W., Seidenberg A. Intuitive probability on finite sets // Annals of mathematical statistics. – 1959. – N 30. – P. 408 – 419.

26. Подиновский Вик.В. К вопросу о продолжении частичных бинарных отношений в моделях принятия решений// Вычислительные системы и вопросы принятия решений. Сборник / Под. ред. Л.Н. Королева, П.С. Краснощекова. – М.: Издательство Московского университета, 1991. – С. 166 – 168.

27. Fishburn P.C. Linear extensions of additive partial orders// Order. - 1997-1998. - N 14. - P. 153 – 169.

28. Шрейдер Ю.А. Равенство, сходство, порядок. – М.: Наука, 1971.

29. Warshall S.A theorem on Boolean matrices// Journal of the Association for computing machinery. - 1962. - N 9. - P. 11 – 12.

30. Подиновский В.В. Многокритериальные задачи с упорядоченными по важности однородными критериями// Автоматика и телемеханика. – 1976. – N 11. – С. 118 – 127.

31. Подиновский В.В., Подиновская О.В. Новые многокритериальные решающие правила в теории важности критериев// Доклады академии наук. – 2013. – N 451(1). – С. 21 – 23.

32. Podinovski V.V. Decision making under uncertainty with unknown utility function and rank-ordered probabilities// European journal of operational research. – 2014. – N 239. – P. 537 – 541.

33. Кривцун И.Л., Подиновский В.В. Аналитическое решающее правило для задачи выбора при порядковой информации о предпочтениях и неопределенных факторах// Информационные технологии моделирования и управления. – 2016. – N 6. – С. 450 – 459.

34. Райфа Г. Анализ решений. Введение в проблему выбора в условиях неопределенности. – М.: Наука, 1977.

35. Нелюбин А.П., Подиновский В.В. Алгоритмическое решающее правило, использующее ординальные коэффициенты важности критериев со шкалой первой порядковой метрики// Журнал вычислительной математики и математической физики. – 2012. – N 1. – С. 43–59.

36. Гафт М.Г., Подиновский В.В. О построении решающих правил в задачах принятия решений// Автоматика и телемеханика. – 1981. – N 6. – С. 128 – 138.

37. Подиновский В.В. Согласительные решения многокритериальных задач выбора// Проблемы управления. – 2017. – N 2. – С. 17 – 26.

38. Подиновский В.В. Количественная важность критериев// Автоматика и телемеханика. – 2000. – N 5. – С. 110 – 123.

39. Podinovski V.V. The quantitative importance of criteria for MCDA// Journal of multi-criteria decision analysis. – 2002. – N 11. – P. 1 – 15.

40. Podinovski V.V. On the use of importance information in MCDA problems with criteria measured on the first ordered metric scale// Journal of multi-criteria decision analysis. – 2009. – N 15. – P. 163 – 174.

41. Подиновский В.В. Анализ решений в условиях неопределенности при нечисловом оценивании предпочтений и вероятностей// Проблемы управления (в печати).

42. Nash J.F. The Bargaining Problem// Econometrica. – 1950. – N 18(2). – P. 155 – 162.

43. Sen A.K. Collective Choice and Social Welfare. – San Francisco: Holden-Day, 1970.

44. Arrow K.J. Social Choice and Individual Values. – New York: Wiley, 1951.

45. Wilson R.B. Social Choice without the Pareto Principle// Journal of Economic Theory. – 1972. – N 5. – P. 478 – 486.

46. d'Aspremont C., Gevers L. Equity and the informational basis of collective choice// Review of Economic Studies. – 1977. – N 44. – P. 199 – 209.

47. May K.O. A set of independent necessary and sufficient conditions for simple majority decisions// Econometrica. – 1952. – N 20(4). – P. 680 – 684.

48. Condorcet. Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. – Paris: L'imprimerie royale, 1785.

49. Black D. On the rationale of group decision-making// Journal of Political Economy. – 1948. – N 56. – P. 23 – 34.

50. Copeland A.H. A reasonable social welfare function (mimeo). – University of Michigan, Ann Arbor (Seminar on Application of Mathematics to the Social Sciences), 1951.

51. Rubinstein A. Ranking the Participants in a Tournament// SIAM Journal on Applied Mathematics. – 1980. – N 38(1). – P. 108 – 111.

52. Fishburn P.C. Condorcet Social Choice Functions// SIAM Journal on Applied Mathematics. – 1977. – N 33(3). – P. 469 - 489.

53. Miller N. A new solution set for tournaments and majority voting: Further graph-theoretical approaches to the theory of voting// American Journal of Political Science. – 1980. – N 24. – P. 68 – 96.

54. Dutta B. Covering sets and a New Condorcet Choice Correspondence// Journal of Economic Theory. – 1988. – N 44. – P. 63 – 80.

55. Laffond G., Laslier J.-F., Le Breton, M.L. The Bipartisan Set of a Tournament Game// Games and Economic Behavior. – 1993. – N 5(1). – P. 182 - 201.

56. Schwartz T. Rationality and the Myth of the Maximum// Noûs. – 1972. – N 6. – P. 97 – 117.

57. Schwartz T. Collective choice, separation of issues and vote trading// The American Political Science Review. – 1977. – N 71(3). – P. 999 – 1010.

58. Good I.J. A note on Condorcet sets// Public Choice. – 1971. – N 10. – P. 97 – 101.

59. Smith J. Aggregation of Preferences with Variable Electorates// Econometrica. – 1973. – N 41(6). – P. 1027 – 1041.

60. Subochev A. Dominant, Weakly Stable, Uncovered Sets: Properties and Extensions// Working paper WP7/2008/03. Moscow: SU – Higher School of Economics. 2008.

61. Inada K. A Note on the Simple Majority Decision Rule// Econometrica. – 1964. – N 32(4). – P. 525 – 531.

62. Inada K. The Simple Majority Decision Rule// Econometrica. - 1969. - N 37(3). – P. 490-506.

63. Sen A.K. A Possibility Theorem on Majority Decisions// Econometrica. – 1966. – N 34(2). – P. 491 – 499.

64. Elkind E., Lackner M., Peters D.. Structured preferences// In U. Endriss, ed., Trends in Computational Social Choice. – 2017.

65. Faliszewski P., Hemaspaandra E., Hemaspaandra L.A., Rothe J. The shield that never was: Societies with single-peaked preferences are more open to manipulation and control// Inform. Comput. – 2011. – N 209(2). – P. 89 – 107.

66. Bredereck R., Chen J., Woeginger G.J. Are there any nicely structured preference profiles nearby?// Math. Soc. Sci. – 2016. – N 79. – P. 61 – 73.

67. Faliszewski P., Hemaspaandra E., Hemaspaandra L.A. The complexity of manipulative attacks in nearly single-peaked electorates// Artif. Intell. - 2014. – N 207. – P. 69 – 99.

68. Elkind E., Faliszewski P., Slinko A. Clone structures in voters' preferences. In Proceedings of the 13th ACM Conference on Electronic Commerce (EC). – 2012. – P. 496 – 513.

69. Ballester M.A., Haeringer G. A characterization of the single-peaked domain// Soc. Choice Welf. – 2011. – N 36. – P. 305 – 322.

70. West J. Generating trees and the Catalan and Schröder numbers// Discrete Math. – 1995. – N 146. – P. 247 – 262.

71. Stankova Z.E. Forbidden subsequences// Discrete Math. – 1994. – N 132. – P. 291 – 316.

72. Ehrenfeucht A., Harju T., ten Pas P., Rozenberg G. Permutations, parenthesis words, and Schröder numbers// Discrete Math. – 1998. – N 190. – P. 259 – 264.

73. West J. Generating trees and forbidden subsequences// Discrete Math. - 1996. – N 157. – P. 363 – 374.

74. Bose P., Buss J.F., Lubiw A. Pattern matching for permutations// Inf. Process. Lett. – 1998. – N 65. - P. 277 - 283.

75. Albert M., Homberger C., Pantone J. Equipopularity classes in the separable permutations// Electron. J. Comb. – 2015. – N 22(2). – P. 2.

76. Kitaev S. Why such patterns? A few motivation points. In Patterns in Permutations and Words. Springer-Verlag Berlin Heidelberg. – 2011. - P. 29 – 80.

77. Gehrlein W.V., Fishburn P.C. The probability of the paradox of voting: a computable solution// J. Econ. Theory. – 1976. – N 13. – P. 14-25.

78. Gehrlein W.V, Fishburn P.C. Proportions of profiles with a majority candidate// Comput. Math. Appl. – 1979. – N 5. – P. 117 - 124.

79. Lackner M.L., Lackner M. On the likelihood of single-peaked preferences// Soc. Choice Welf. - 2017. – N 48(4). – P. 717 – 745.

80. Chen J., Finnendahl U.P. On the number of single-peaked narcissistic or single-crossing narcissistic preference profiles// Discrete Math. – 2018. – N 341. – P. 1225 – 1236.

81. Monjardet B. Acyclic domains of linear orders: a survey. In: Brams, S., Gehrlein, W., Roberts, F. (eds.) The Mathematics of Preference, Choice and Order. – 2009. – P. 136–160.

82. Danilov V.I., Koshevoy G.A. Maximal Condorcet Domains// Order. - 2013. - N 30. – P. 181 – 194.

83. Puppe C. The single-peaked domain revisited: A simple global characterization// J. Econ. Theory. – 2018. – N 176. – P. 55 – 80.

84. Murtagh F. Counting dendrograms: a survey// Discrete Appl. Math. – 1984. – N 7. – P. 191 - 199.

85. Karpov A. On the number of group-separable preference profiles// Group Decis. Negot. – 2019. – N 28(3). – P. 501 – 517.

86. Gale D., Shapley L. College admission and the stability of marriage// Am. Math. Monthly. – 1962. – N 69. – P. 9 – 15.

87. Eeckhout J. On the uniqueness of stable marriage matchings// Econom. Lett. – 2000. – N 69(1). – P. 1 – 8.

88. Clark S. The uniqueness of stable matchings// Contrib. Theor. Econ. – 2006. – N 6(1). - Art. 8.

89. Niederle M., Yariv L. Decentralized matching with aligned preferences// NBER Working Paper. – 2009. – N 14840.

90. Holzman R., Samet, D. Matching of Like Rank and the Size of the Core in the Marriage Problem// Games Econ. Behav. – 2014. – N 88. – P. 277 – 285.

91. Romero-Medina A., Triossi M. Acyclicity and singleton cores in matching markets// Econom. Lett. – 2013. – N 118(1). – P. 237 – 239.

92. Park J. Competitive equilibrium and singleton cores in generalized matching problems// Int. J. Game Theory. – 2017. – N 46(2). – P. 487 – 509.

93. Ramani V., Mallikarjuna Rao K.S. Paths to stability and uniqueness in two-sided matching markets// Int. J. Game Theory. – 2018. – N 47(4). – P. 1137 - 1150.

94. Wu Q. A finite decentralized marriage market with bilateral search// J. Econ. Theory. – 2015. – N 160. – P. 216 – 242.

95. Karpov A. A necessary and sufficient condition for uniqueness consistency in the stable marriage matching problem// Econom. Lett. – 2019. – N 178. – P. 63 – 65.

96. Kesten O. School choice with consent// Q. J. Econ. – 2010. – N 125(3). – P. 1297 – 1348.

97. Thomson W. Consistency and its converse: an introduction// Rev. Econ. Design. – 2011. – N 15(4). – P. 257 – 291.

98. Klaus B. Consistency and its converse for roommate markets// Games Econ. Behav. 2017. – N 104. – P. 43 – 58.

99. Kivelä M., Arenas A., Barthelemy M., Gleeson J.P., Moreno Y., Porter M.A. Multilayer networks// Journal of Complex Networks. – 2014. – N 2(3). – P. 203 – 271.

100. Wasserman S., Faust K. Social Network Analysis: Methods and Applications. – Cambridge: Cambridge University Press. – 1994.

101. Bianconi G. Statistical mechanics of multiplex networks: entropy and overlap// Phys. Rev. E. – 2013. – N 87. – 062806.

102. Horvát E.Á., Zweig K.A. One-mode Projection of Multiplex Bipartite Graphs// IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. – 2012. – Istanbul.

103. Rocklin M., Pinar A. On clustering on graphs with multiple edge types// Internet Math. – 2013. – N 9(1). – P. 82 - 112.

104. Cozzo E., de Arruda G.F., Rodrigues F.A., Moreno Y. Multilayer Networks: Metrics and Spectral Properties// Interconnected Networks. Understanding Complex Systems. Springer, Cham. – 2016.

105. Solé-Ribalta A., De Domenico M. Kouvaris N.E., Díaz-Guilera A., Gómez S., Arenas A. Spectral properties of the Laplacian of multiplex networks// Phys. Rev. E. – 2013. – N 88. – 032807.

106. De Domenico M., Sole-Ribalta A., Cozzo E., Kivela M., Moreno Y., Porter M. A., Gomez S., Arenas A. Mathematical formulation of multi-layer networks// Physical Review X. – 2013. – N 3. – 041022.

107. Rahmede C., Iacovacci J., Arenas A., Bianconi G. Centralities of nodes and influences of layers in large multiplex networks//Journal of Complex Networks. – 2017. – cnx050.

108. Solá L., Romance M., Criado R., Flores J., García del Amo A., Boccaletti S. Eigenvector centrality of nodes in multiplex networks// Chaos. – N 23. – Article 033131.

109. De Domenico M., Sole-Ribalta A., Omodei E., Gomez S., Arenas A. Centrality in interconnected multilayer networks. // arXiv preprint arXiv:1311.2906. – 2013.

110. Girvan M., Newman M. E. J. Community structure in social and biological networks// Proceedings of the National Academy of Sciences. – 2002. – N 99(12). – P. 7821 - 7826;

111. De Meo P., Ferrara E., Fiumara G, Ricciardello A. A Novel Measure of Edge Centrality in Social Networks// Knowledge-Based Systems. – 2012. – 30. 10.1016/j.knosys.2012.01.007.

112. Page L.B., Perry J.E. Reliability polynomials and link importance in networks// IEEE Transactions on Reliability. – 1994. – N 43(1). – P. 51 – 58.

113. Nicholson C.D., Barker K., Ramirez-Marquez J.E. Flow-based vulnerability measures for network component importance: Experimentation with preparedness planning// Reliability Engineering System Safety. – 2016. – N 145. – P. 62 – 73.

114. Mavroforakis C., Garcia-Lebron R., Koutis I., Terzi E. Spanning Edge Centrality: Large-scale Computation and Applications// WWW. – 2015.

115. Acemoglu D., Carvalho V.M., Ozdaglar A., Tahbaz-Salehi A. The network origins of aggregate fluctuations// Econometrica. – 2012. – N 80(5). – P. 1977 – 2016.

116. Acemoglu D., Ozdaglar A., Tahbaz-Salehi A. Networks, shocks, and systemic risk// Tech. rep. National Bureau of Economic Research. – 2015 – N w20931.

117. Aleskerov F., Manfred J. H., Kamalova R. Power distribution in the Weimar Reichstag in 1919–1933//Annals of Operations Research. – 2014. – N 215(1). – P. 25 - 37.

118. Latora V., Massimo M. How the science of complex networks can help developing strategies against terrorism// Chaos, solitons & fractals. – 2004. – N 20(1). – P. 69 – 75.

119. Kuikka V. Terrorist Network Analyzed with an Influence Spreading Model// International Workshop on Complex Networks. – Springer, Cham, 2018. – P. 185 – 197.

120. Latora V., Massimo M. Efficient behavior of small-world networks// Physical review letters. – 2001. – N 87.19. – P.198701.

121. Von Ferber C., Holovatch T., Holovatch Y., Palchykov V. Public transport networks: empirical analysis and modeling// The European Physical Journal B. – 2009. – N 68(2). – P. 261 – 275.

122. Felsenthal D.S., Mosh́e M. Voting power measurement: a story of misreinvention// Social choice and welfare. – 2005. – 25. – N 2-3. – P. 485 – 506.

123. International Migrant Stock 2019. [online] Available at: https://www.un.org/en/development/desa/population/migration/publications/migrationreport/docs/MigrationStock2019_TenKeyFindings.pdf [Accessed 1 Nov. 2019].

124. Fagiolo G. Mastrorillo M. International migration network: Topology and modeling// Physical Review E. – 2013. – 88(1).

125. Aleskerov F.T., Meshcheryakova N.G., Rezyapova A., Shvydun S. Network Analysis of International Migration// Models, Algorithms, and Technologies for Network Analysis. Springer Proceedings in Mathematics & Statistics. – 2017. – N 197. – P. 177 – 185.

126. Lovric M., Da Re R., Vidale E., Pettenella D., Mavsara R. Social network analysis as a tool for the analysis of international trade of wood and non-wood forest products// Forest Policy and Economics. – 2018. – N 86. – P. 45 – 66.

127. Aleskerov F., Meshcheryakova N., Shvydun S. Power in Network Structures// Springer, Cham. - Springer Proceedings in Mathematics & Statistics. – 2017. – Models, Algorithms, and Technologies for Network Analysis. – N 197. – P. 79 – 85.

128. Aleskerov F., Meshcheryakova N., Shvydun S. Centrality measures in networks based on nodes attributes, long-range interactions and group influence// arXiv preprint arXiv:1610.05892. – 2016.

129. Meshcheryakova N., Shvydun S. Power in Network Structures Based on Simulations// Studies in Computational Intelligence. – 2018. – N 689: Complex Networks & Their Applications VI. – Springer. – P. 1028 – 1038.

130. Aizerman M., Aleskerov F. Theory of Choice. – Elsevier. – North-Holland. – 1995.

131. Shvydun S. Normative properties of multi-criteria choice procedures and their superpositions: I// Working paper WP7/2015/07 (Part 1). – 2015. – Moscow: HSE Publishing House.

132. Aleskerov F.T., Subochev A. Matrix-vector representation of various solution concepts// Working papers WP7/2009/03. – 2009. – Moscow: HSE Publishing House.

133. Aleskerov F., Sergeeva Z., Shvydun S. Assessment of Exporting Economies Influence on the Global Food Network// Optimization Methods and Applications. Springer Optimization and Its Applications. – 2017. – N 130. – Springer, Cham. – P. 1 – 10.

134. Aleskerov F.T., Meshcheryakova N.G., Sergeeva Z., Shvydun S.V. Centrality Measures And Clustering Analysis in a Retail Food Network// 2017 IEEE 11th International Conference on Application of Information and Communication Technologies. – 2017. – N 1. – M.: Institute of Electrical and Electoronics Engineers. – P. 48 – 52.

135. SIPRI. The SIPRI Arms transfers database. Accessed: 2019-07-01.

136. Bonacich P. Power and centrality: A family of measures// American Journal of Sociology. – 1987. – N 92(5). – P. 1170 – 1182.

137. Shapley L.S., Shubik M. A method for Evaluating the Distribution of Power in a Committee System// American Political Science Review. – 1954. – N 48(3). – P. 787 – 792.

138. Myerson R.B. Graphs and cooperation in games// Mathematics of Operations Research. – 1977. – N 2(3). – P. 225 – 229.

139. Khmelnitskaya A., Selçuk Ö., Talman D. The Shapley value for directed graph games// Operations Research Letters. – 2016. – N 44(1). – P. 143 – 147.

140. Skibski O., Rahwan T., Michalak T. P., Yokoo M. Attachment centrality: An axiomatic approach to connectivity in networks// Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems. – International Foundation for Autonomous Agents and Multiagent Systems, 2016. – P. 168 – 176.

141. Dubey P., Shapley L.S. Mathematical properties of the Banzhaf power index// Mathematics of Operations Research. – 1979. – N 4(2). – P. 99 – 131.

142. Aleskerov F. Power indices taking into account agents' preferences// Mathematics and democracy. – 2006 – Springer, Berlin, Heidelberg. – P. 1 – 18.

143. Boldi P., Vigna S. Axioms for centrality// Internet Mathematics. – 2014. – N 10(3-4). – P. 222 – 262.

144. Gomez D., González-Arangüena E., Manuel C., Owen G., del Pozo M., Tejada J. Centrality and power in social networks: a game theoretic approach// Mathematical Social Sciences. – 2003. – N 46(1). – P. 27 – 54.

145. Shvarts D. A. Axiomatics for power indices that account for player preferences// Automation and Remote Control. – 2010. – N 71(1). – P. 128 – 141.

146. Laruelle A., Valenciano F. Shapley-Shubik and Banzhaf indices revisited// Mathematics of Operations Research. – 2001. – N 26(1). – P. 89 – 104.

147. Jackson M. O. Social and economic networks. – Princeton: Princeton University Press. – 2010.

148. International Migrant Stock 2015. [online] Available at: https://www.un.org/en/development/desa/population/migration/publications/migrationreport/docs/MigrationStock2019_TenKeyFindings.pdf [Accessed 1 April 2016].

149. Migration and migrant population statistics - Statistics Explained. [online] Available at: http://ec.europa.eu/eurostat/statistics-explained/index.php/Migration_and_migrant_population_statistics [Accessed 4 Oct. 2019].

150. International Migration Database. [online] Available at: http://stats.oecd.org/Index.aspx?DataSetCode=MIG [Accessed 3 Oct. 2019].

151. Hilgerdt F. The Case for Multilateral Trade// The American Economic Review. – 1943. – N 33(1). – P. 393 – 407.

152. Benedictis L., Tajoli L. The world trade network// The World Economy. – 2011. – N 34(8). – P. 1418 – 1455.

153. Benedictis L., Nenci S., Santoni G., Tajoli L., Vicarelli C. Network analysis of world trade using the BACI-CEPII dataset// CEPII Working Paper. – 2013. – N 14(3-4). – P. 60.

154. Zhang S., Wang L., Liu Z., Wang X. Evolution of international trade and investment networks// Physica A. – 2016. – N 462. – P. 752 – 763.

155. Meshcheryakova N.G. The Impact of Indirect Connections: The Case of Food Security Problem// Studies in Computational Intelligence. 2019. – N 813. Complex Networks and Their Applications VII. Part 2. – Springer. – Ch. 7. – P. 80 – 90.

156. Song Y., Liu S., Wang H., Wang Z., Li H. Automatic taxonomy construction from keywords// U.S. Patent No. 9,501,569. Washington, DC: U.S. Patent and Trademark Office, 2016.

157. Wang C., He X., Zhou A. A Short Survey on Taxonomy Learning from Text Corpora: Issues, Resources and Recent Advances// Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. – 2017. – P. 1190 - 1203.

158. Vedula N., Nicholson P.K., Ajwani D., Dutta S., Sala A., Parthasarathy S. Enriching Taxonomies With Functional Domain Knowledge// The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, ACM. – 2018. P. 745 - 754.

159. Lloret E., Boldrini E., Vodolazova T., Martanez-Barco P., Munoz R., Palomar M. A novel concept-level approach for ultra-concise opinion summarization// Expert Systems with Applications, N 42(20). – 2015. P. 7148 - 7156.

160. Mueller G, Bergmann R. Generalization of Workflows in Process-Oriented Case-Based Reasoning// FLAIRS Conference. – 2015. – P. 391 - 396.

161. Waitelonis J., Exeler C., Sack H. Linked data enabled generalized vector space model to improve document retrieval// Proceedings of NLP & DBpedia 2015 workshop in conjunction with 14th International Semantic Web Conference (ISWC), CEUR-WS, N 1486, 2015.

162. Mirkin B., Orlov M. Three aspects of the research impact by a scientist: measurement methods and an empirical evaluation// Optimization, Control, and Applications in the Information Age, Springer, Cham. – 2015. – P. 233 - 259.

163. Frolov D., Mirkin B., Nascimento S., Fenner T. Finding an appropriate generalization for a fuzzy thematic set in taxonomy// Working paper WP7/2019/01, Moscow, Higher School of Economics Publ. House. – 2018. – 60p.

164. Blei D. Probabilistic topic models// Communications of the ACM. – N 55(4). – 2012. – P. 77 - 84.

165. Nascimento S., Fenner T., Mirkin B. Representing research activities in a hierarchical ontology// Procs. of 3rd International Workshop on Combinations of

Intelligent Methods and Applications (CIMA 2012), Montpellier, France, August 28. – 2012. – P. 23 - 29.

166. Gene ontology consortium: going forward// Nucleic Acids Research, vol. 43 D1049-D1056, 2015.

167. Lee D., Cornet R., Lau F., De Keizer N. A survey of SNOMED CT implementations// Journal of Biomedical Informatics. – N 46(1). – 2013. – P. 87 - 96.

168. Chernyak E. An approach to the problem of annotation of research publications// Proceedings of the eighth ACM international conference on web search and data mining, ACM. -– 2015. – P. 429 - 434.

169. Mirkin B., Nascimento S. Additive spectral method for fuzzy cluster analysis of similarity data including community structure and affinity matrices// Information Sciences. – 2012. – N 183(1). – P.16-34.

170. Frolov D., Nascimento S., Fenner T., Mirkin B. Parsimonious Generalization of Fuzzy Thematic Sets in Taxonomies Applied to the Analysis of Tendencies of Research in Data Science// Information Sciences. 2019.

171. Frolov D., Mirkin B., Nascimento S., Fenner T. Method for Generalization of Fuzzy Sets// Proceedings of the 18th International Conference on Artificial Intelligence and Soft Computing (ICAISC-2019), Springer Lecture Notes in Artificial Intelligence (LNAI). – 2019. – P. 273 - 286.

172. Frolov D., Mirkin B., Nascimento S., Fenner T. Using Taxonomy Tree to Generalize a Fuzzy Thematic Clusters// IEEE 2019 International Conference on Fuzzy Systems Proceedings. IEEE. – 2019.

173. Frolov D., Nascimento S., Fenner T., Taran Z., Mirkin B. Computational Generalization in Taxonomies Applied to: (1) Analyze Tendencies of Research and (2) Extend User Audiences// IDEAL 2019 Proceedings. Springer. – 2020.

174. Frolov D., Mirkin B., Nascimento S., Fenner T. Using Domain Taxonomy to Model Generalization of Thematic Fuzzy Clusters// IARIA Content 2019 Proceedings. IARIA. – 2019. – P. 20 - 25.

175. Frolov D., Mirkin B., Nascimento S., Fenner T. Globally Optimal Parsimoniously Lifting a Fuzzy Query Set Over a Taxonomy Tree// Proceedings of the World Congress on Global Optimization, Springer, LNCS. – 2019. – P. 781 - 789.

176. Inselberg A. The plane with parallel coordinates// The Visual Computer. – 1985. – N 1(2). – P. 69 - 91.

177. Few S. Multivariate analysis using parallel coordinates// Perceptual Edge. – 2006. – P. 1 - 9.

178. Aleskerov F., Alper C. A clustering approach to some monetary facts: a long-run analysis of cross-country data// The Japanese Economic Review. – 2000. – N 51(4). – P. 555 – 567.

179. Aleskerov F., Ersel H., Yolalan R. Multicriterial Ranking Approach for Evaluating Bank Branch Performance// International Journal of Information Technology and Decision Making. – 2004. – N 3(2). – P. 321 – 335.

180. Aleskerov F., Nurmi H. A method for finding patterns of party support and electoral change: An analysis of British general and Finnish municipal elections// Mathematical and Computer Modelling. – 2008. – N 48. – P. 1385 – 1395.

181. Егорова Л.Г., Мячин А.Л. Структура российского сообщества экономистов и его отношение к российским экономическим журналам. Ч.2. Анализ паттернов респондентов// Проблемы управления. – 2019. – N 4. – P. 50 - 57.

182. Makwana H., Tanwani S., Jain S. Axes Re-ordering in Parallel Coordinate for Pattern Optimization// International Journal of Computer Applications. – 2012. – 40(13).

183. Myachin A. New methods of pattern analysis in the study of Iris Anderson-Fisher Data// In: Computers Communications and Control (ICCCC), 2016 6th International Conference, IEEE. 2016. – P. 97 - 102.

184. Мячин А.Л. Анализ паттернов в системе параллельных координат на базе парного сравнения показателей // Автоматика и телемеханика. – 2019. – N 1. – С. 138 - 152.

185. Мячин А.Л. Определение центроидов для повышения точности порядково-инвариантной паттерн-кластеризации// Управление большими системами. – 2019. – N 78. – С. 6 - 22.

186. Myachin A. A Symmetry of the Data Pattern Structure (in press).

187. Siegler R.S. Three Aspects of Cognitive Development// Cognitive Psychology. –1976. – N 8(4). – P. 481 - 520.

188. Мячин А.Л. Новые методы анализа паттернов на примере набора данных Balance Scale Data Set// Управление большими системами. – 2016. – С. 134 – 142.

189. Peng W., Ward M., Rundensteiner E. Clutter reduction in multi-dimensional data visualization using dimension reordering. In Proceedings of the Symposium on Information Visualization, P. 89 – 96, Los Alamitos, USA, 2004. IEEE Computer Society.

190. Wegman E.J., Luo Q. High dimensional clustering using parallel coordinates and the grand tour// Classification and Knowledge Organization. – Springer, Berlin, Heidelberg, 1997. – P. 93 - 101.

191. Artero A., de Oliveira M., Levkowitz H. Uncovering Clusters in Crowded Parallel Coordinates Visualizations. In Proc. of IEEE Symposium on Information Visualization, P. 81 - 88, Austin, Texas, USA, 2004. IEEE Computer Society.

192. Brandenburger A. M., Nalebuff B. J. Co-Opetition. – Crown Business, 2011. – 304 с.

193. Korgin N., Østbye S. Coopetition in the Arctic: Theory and Experiments// Proceedings of the 18th International Conference on Group Decision and Negotiation. NUAA Discussion Papers in Economics and Management. – 2018. – P. 356-369.

194. Belenky A., Fedin G., Kornhauser A. Estimating the needed volume of investment in a public-private partnership to develop a regional energy/transportation infrastructure// International Journal of Public Administration. – 2019. – N 42(15-16). – P. 1275 – 1310.

195. Belenky A., Farazmand A., Vasin A. The management of large scale energy projects: opportunities and challenges// International Journal of Public Administration. – 2019. – N 42(15-16). – P. 1251 – 1255.

196. Belenky A., Kornhauser A. Estimating the needed volume of investment in developing a chain of charging stations for electric vehicles along a highway// International Journal of Public Administration. – 2019. – N 42(15-16). – P. 1256 – 1274.

197. Fallahi F. Stochastic convergence in per capita energy use in world// Energy Economics. – 2017. – N 65. – P. 228 – 239.

198. Coviello M., Gollán J., Pérez M. Public-private partnerships in renewable energy in Latin America and the Caribbean// Retrieved from repositorio.cepal.org. – 2012.

199. Lenferna G. Can we equitably manage the end of the fossil fuel era?// Energy Research & Social Science. – 2018. – N 35.– P. 217 – 223.

200. Barreto R. Fossil fuels, alternative energy and economic growth// Economic Modelling. – 2018. – N 75. – P. 196 – 220.

201. Speirs J., McGlade C., Slade R. Uncertainty in the availability of natural resources: Fossil fuels, critical metals and biomass// Energy Policy. – 2015. – N 87. – P. 654 – 664.

202. Evans A., Strezov V., Evans T. Sustainability concepts of energy generation technologies/ Encyclopedia of Sustainable Technologies. – 2017. – P. 3 – 10.

203. Kis Z., Pandya N., Koppelaar R. Electricity generation technologies: Comparison of materials use, energy return on investment, jobs creation and CO2 emissions reduction// Energy Policy. – 2018. – N 12. – P. 144 – 157.

204. Feng Z., Niu W., Cheng C. Optimal allocation of hydropower and hybrid electricity injected from inter-regional transmission lines among multiple receiving-end power grids in China// Energy. – 2018. – N 162. – P. 444 – 452.

205. Kumar P., Ali I., Thomas M. Energy efficiency analysis of reconfigured distribution system for practical loads// Perspectives in Science. – 2016. – N 8. – P. 498 – 501.

206. Gordon D., Gunsch K., Pawluk C. A natural monopoly in natural gas transmission// Energy Economics. – 2003. – N 25(5). – P. 473 – 485.

207. Sørensen B. Energy transmission and storage// In Renewable energy. – 2017 (5th ed.). – P. 569 – 646.

208. Mete E., Turkay M. Energy network optimization in an oil refinery// Computer Aided Chemical Engineering. – 2018. – N 44. – P. 1897 – 1902.

209. Hollermann D., Hoffrogge D., Mayer F., Hennen M., Bardow A. Optimal (n−1)-reliable design of distributed energy supply systems// Computers & Chemical Engineering. – 2019. – N 121(2). – P. 317 – 326.

210. Han S., Kim J. An optimization model to design and analysis of renewable energy supply strategies for residential sector// Renewable Energy. – 2017. – N 112. – P. 222 – 234.

211. Belenky A., Larson R., Roginskiy L. Storing electricity in a country's electrical grid as a key energy problem of the 21st century// Procedia Computer Science. – 2016. – N 91. – P. 873 – 881.

212. Zhang X., Lovati M., Vigna I., Widén J., Feng T. A review of urban energy systems at building cluster level solutions// Applied Energy. – 2018. – N 230(15). – P. 1034 –1056.

213. Seefeldt P. A stowing and deployment strategy for large membrane space systems on the example of Gossamer-1// Advances in Space Research. – 2017. – N 60(6). – P. 1345 – 1362.

214. Das C., Bass O., Kothapalli G., Mahmoud T., Habibi D. Overview of energy storage systems in distribution networks: Placement, sizing, operation, and power quality// Renewable and Sustainable Energy Reviews. – 2018. – N 91. – P. 1205 – 1230.

215. Zamani-Dehkordi P., Shafiee S., Rakai L., Knight A., Zareipour H. Price impact assessment for large-scale merchant energy storage facilities// Energy. – 2017. – N 125. – P. 27 – 43.

216. Demski C., Thomas G., Becker S., Evensen D., Pidgeon N. Acceptance of energy transitions and policies: Public conceptualisations of energy as a need and basic right in the United Kingdom// Energy Research & Social Science. – 2019. – N 48. – P. 33 – 45.

217. Chen H.-W., Yu R.-F., Chuang Y.-H. Bio-inspired optimal site selection of LPG stations for gas-driven cars in an urban region// Journal of Natural Gas Science and Engineering. – 2016. – N 35(A). – P. 1301 – 1309.

218. Отчет МНУЛ анализа и выбора решений, ТЗ-72 за 2017г.

219. The German Federal Agency for Civic Education// http://www.bpb.de/politik/ (date of access: 18.01.2019)

220. Tangian A. Mathematical Theory of Democracy. – Springer-Verlag Berlin Heidelberg. – 2014.

221. Harmel R., Uk H., Tan A., Janda K. Performance, Leadership, Factions, and Party Change: An Empirical Analysis// West European Politics. – 1995. – N 18(1). – P. 1 – 33.

222. Bernauer J., Brauninger T. Intra-Party Preference Heterogeneity and Faction Membership in the 15th German Bundestag: A Computational Text Analysis of Parliamentary Speeches// German Politics. – 2008. – N 18(3). – P. 385 – 402.

223. Schofield N. Stability of coalition governments in Western Europe: 1945-1986// European Journal of Political Economy. – 1987. – N 3(4). – P. 555 – 591.

224. Levy G. A Model of Political Parties// Journal of Economic Theory. – 2004. – N 115(2). – P. 250 – 277.

225. Landa D., Meirowitz A. Game Theory, Information, and Deliberative Democracy// American Journal of Political Science. – 2009. – N 53(2). – P. 427 – 444.

226. Persico N., Rodriguez-Pueblita J., D. Silverman D. Factions and Political Competition// Journal of Political Economy. – 2011. – N 2. – P. 242 – 288.

227. Савватеев А.В. Анализ коалиционной устойчивости «биполярного мира»// Журнал Новой экономической ассоциации. – 2013. – N 1(17). – С. 10 – 43.

228. Aleskerov F., Ersel H., Sabuncu Y. Power and coalitional stability in the Turkish Parliament (1991-1999)// Turkish Studies. – 2000. – N 1(2). – P. 21 – 38.

229. Алескеров Ф.Т., Соколова А.В., Благовещенский Н.Ю., Сатаров Г.А., Якуба В.И. Влияние и структурная устойчивость в российском парламенте (1905-1917 и 1993-2005 гг.). – М.: Физматлит. – 2007.

230. Grabisch M. Set Functions, Games and Capacities in Decision Making. – Cham: Springer. – 2016.

231. Dempster A.P. Upper and lower probabilities induced by multivalued mapping// Ann. Math. Statist. – 1967. – N 38. – P. 325 – 339.

232. Shafer G. A mathematical theory of evidence. – Princeton, N.J.: Princeton University Press. – 1976.

233. The German Bundestag// http://www.bundestag.de/ (date of access: 18.01.2019).

234. Choquet G. Theory of capacities// Annales de l'Institut Fourier. – 1953. – N 5. – P. 131 – 295.

235. Grabisch M., Nicolas J.-M. Classification by fuzzy integral: performance and tests// Fuzzy Sets and Systems. – 1994. – N 65(2-3). – P. 255 – 271.

236. Keller J.M., Gader P.D., Tahani H., Chiang J.-H., Mohamed M. Advances in fuzzy integration for pattern recognition// Fuzzy Sets and Systems. – 1994. – N 65. – P. 273 – 283.

237. Grabisch M. Fuzzy integral for classification and feature extraction// in Grabisch M., Murofushi T., Sugeno M. (eds.), Fuzzy Measures and Integrals: Theory and Applications. – 2000. – P. 415 – 434.

238. Grabisch M., Roubens M. Application of the Choquet integral in multicriteria decision making// in Grabisch M., Murofushi T., Sugeno M. (eds.), Fuzzy Measures and Integrals: Theory and Applications. – 2000. – P. 348 – 374.

239. Su K.-L., Jau Y.-M., Jeng J.-T. Modeling of Nonlinear Aggregation for Information Fusion Systems with Outliers Based on the Choquet Integral// Sensors. – 2011. – N 11. – P. 2426 – 2446.

240. Du X. Multiple Instance Choquet Integral For MultiResolution Sensor Fusion. – Ph.D Thesis, Columbia, MO. – 2017.

241. Grabisch M. k-order additive discrete fuzzy measures and their representation// Fuzzy Sets and Systems. – 1997. – N 92. – P. 167 – 189.

242. Chateauneuf A., Jaffray J.-Y. Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion// Mathematical Social Sciences. – 1989. – N 17. – P. 263 – 283.

243. Shapley L. A value for n-person games// Annals of Mathematical Studies. – 1953. – N 28. – P. 307 – 317.

244. Smets P. Decision making in TBM: the necessity of the pignistic transformation// International Journal of Approximate Reasoning. – 2005. – N 38. – P. 133 – 147.

245. Bronevich A., Lepskiy A., Penikas H. The Application of Conflict Measure to Estimating Incoherence of Analyst's Forecasts about the Cost of Shares of Russian Companies// Procedia Computer Science. – 2015. – N 55. – P. 1113 – 1122.

246. Destercke S., Burger T. Toward an axiomatic definition of conflict between belief functions// IEEE Transactions on Cybernetics. – 2013. – N 43(2). – P. 585 – 596.

247. Liu W. Analysing the degree of conflict among belief functions// Artificial Intelligence. – 2006. – N 170. – P. 909 – 924.

248. Martin A. About Conflict in the Theory of Belief Functions// Belief Functions: Theory and Applications, Advances in Intelligent and Soft Computting. – 2012. – AISC. – N 164. – P. 161 – 168.

249. Daniel M. Conflicts within and between Belief Functions// In: E. Hüllermeier, R. Kruse, E. Hoffmann (eds.) IPMU 2010. – LNAI 6178. – P. 696 – 705. Springer-Verlag, Berlin Heidelberg. – 2010.

250. Lepskiy A. On internal conflict as an external conflict of a decomposition of evidence// In Vejnarová J., Kratochvil V. (eds.): BELIEF 2016. – LNAI 9861. – Springer-Verlag. – P. 25 – 34.

251. Schubert J. The internal conflict of a belief function// In: T. Denoeux, M-H. Masson (eds.): BELIEF. – 2012. – AISC 164. – P. 169 – 177.

252. Lepskiy A., Smolev V. Application of Non-additive Measures and Integrals for Analysis of the Importance of Party Positions for Voting// in: Atlantis Studies in Uncertainty Modelling. – 2019. – N 1. – Atlantis Press. – P. 321 – 327.

253. Bronevich A., Lepskiy A. Clustering a body of evidence based on conflict measures// in: Atlantis Studies in Uncertainty Modelling. – 2019. – N 1. – Atlantis Press. – P. 328 – 333.

254. Esteban J.M., Ray D. On the measurement of polarization// Econometrica: Journal of the Econometric Society. – 1994. – N 62(4). – P. 819 – 851.

255. Carey J.M. Legislative organization. – The Oxford Handbook of Political Institutions. – 2006. – P. 431 – 451.

256. Shugart M.S., Carey J.M. Presidents and assemblies: Constitutional design and electoral dynamics. – Cambridge University Press. – 1992.

257. Siaroff A. Comparative European party systems: An analysis of parliamentary elections since 1945. – N.Y., L.: Taylor and Francis. – 2000.

258. Голосов Г.В. Партийные системы стран мира: Региональные и хронологическое распределение, модели устойчивости// Политическая наука. – 2012. – N 3. – С. 71 – 104.

259. Laakso M., Taagepera R. Effective» number of parties: A measure with application to West Europe// Comparative Political Studies. – 1979. – N 12. – P. 3 – 27.

260. Penrose L.S. The elementary statistics of majority voting// Journal of the Royal Statistical Society. – 1946. – N 109. – P. 53 – 57.

261. Banzhaf J.F. Weighted Voting Doesn't Work: A Mathematical Analysis// Rutgers Law Review. – 1965. – N 19. – P. 317 – 343.

262. Banzhaf J.F. One man, 3.312 votes: a mathematical analysis of the Electoral College// Villanova Law Review. – 1968. – N 13. – P. 304 – 332.

263. Banzhaf J.F. Multi-member electoral districts – do they violate the "one man, one vote" principle// Yale Law Journal. – 1966. – N 75. – P. 1309 – 1338.

264. Felsenthal D.S., Machover M. The measurment of voting power. – Cheltenham: Edward Elgar Publishing. – 1998.

265. Riker W.H. The first power index// Social Choice and Welfare. – 1986. – N 3(4). – P. 293 – 295.

266. Riker W.H. A test of the adequacy of the power index// Behavioral Science. – 1959. – N 4. – P. 120 – 131.

267. Shapley L.S., Owen G. Optimal location of candidates in ideological space// International Journal of Game Theory. – 1989. – N 1. – P. 125 – 142.

268. Алескеров Ф.Т., Очур О.А. Обобщенные индексы Шепли-Оуэна и распределение влияния в Государственной Думе III созыва// Препринт WP7/2007/03 – М.: ГУ ВШЭ. – 2007.

269. McCoy J., Tahmina R. Polarized Democracies in Comparative Perspective: Toward a Conceptual Framework// International Political Science Association Conference. – 2016.

270. Montalvo J.G., Reynal-Querol M. Why ethnic fractionalization? Polarization, ethnic conflict and growth// Economics Working Papers, Universitat Pompeu Fabra. – N 660. – 2002.

271. Esteban J., Gradín C., Ray D. An extension of a measure of polarization, with an application to the income distribution of five OECD countries// Journal of Economic Inequality. – 2007. – N 5(1). – P. 1 – 19.

272. Wolfson M.C. When inequalities diverge// The American Economic Review. – 1994. – N 84(2). – P. 353 – 358.

273. Wolfson M.C. Divergent inequalities: theory and empirical results// Review of Income and Wealth. – 1997. – N 43(4). – P. 401 – 421.

274. Chakravarty S.R., Majumder A. Inequality, polarisation and welfare: Theory and applications// Australian Economic Papers. – 2001. – N 40(1). – P. 1 – 13.

275. Rodríguez J.G., Salas R. Extended bi-polarization and inequality measures// in Inequality, Welfare and Poverty: Theory and Measurement. – Emerald Group Publishing Limited. – 2003. – P. 69 – 83.

276. Chakravarty S.R., D'Ambrosio C. Polarization orderings of income distributions// Review of Income and Wealth. – 2010. – N 56(1). – P. 47 – 64.

277. Gigliarano C., Mosler K. Constructing indices of multivariate polarization// Journal of Economic Inequality. – 2009. – N 7(4). – P. 435 – 460.

278. Sartori G. Parties and political systems: A framework for analysis. – ECPR Press. – 2005.

279. Downs A. An Economic Theory of Democracy. – New York: Harper & Brothers. – 1957.

280. Vegetti F. Political Polarization in Hungary// Memo prepared for Polarized Politics Workshop, Georgia State University, March 14-15, 2016. – Atlanta, GA. – 2016.

281. Dahl R.A. The American Oppositions: Affirmation and Denial// in R.A. Dahl (ed.), Political Oppositions in Western Democracies, New Haven: Yale University Press. – 1966. – P. 34 – 69.

282. Benjamin N., Cathcart A., Ryan K. Low Default Portfolios: a Proposal for Conservative Estimation of Default Probabilities// http://www.fsa.gov.uk/pubs/international/default_probabilities.pdf. – 2006.

283. Vasicek O. Probability of Loss on Loan Portfolio// http://mx.nthu.edu.tw/~jtyang/Teaching/Risk_management/Papers/Models/Probability%20of%20Loss%20on%20Loan%20Portfolio.pdf. – 1987.

284. Witzany J. A note on the Vasicek's model with the logistic distribution». Ekonomický časopis. – 2013. – N 10. – P. 1053 – 1066.

285. Ермолова М.Д., Пеникас Г.И. Исследование взаимосвязи параметров моделей внутренних рейтингов оценки кредитного риска — вероятности

дефолта и доли убытка при дефолте (часть 1)// Управление финансовыми рисками. – 2015. – N 41(1). – С. 22–45.

286. Ermolova M., Penikas H. The impact of PD-LGD correlation on bank capital adequacy in nongranular loan portfolio// Model Assisted Statistics and Applications. – 2019. – N 14. – P. 1 – 18.

287. Studies on the Validation on Internal Rating Systems// https://www.bis.org/publ/bcbs_wp14.pdf.

288. An Explanatory Note on the Basel II IRB Risk Weight Functions// https://www.bis.org/bcbs/irbriskweight.pdf.

289. International Convergence of Capital Measurement and Capital Standards. A Revised Framework. Comprehensive Version// https://www.bis.org/publ/bcBS128.PDF

290. Положение Банка России от 6 августа 2015 г. №483-П «О порядке расчета величины кредитного риска на основе внутренних рейтингов»// https://base.garant.ru/71203444.

291. Ермолова М.Д., Пеникас Г.И. Биномиальный тест для коррелированных бинарных случайных величин для проверки точности рейтинговой модели// Управление финансовыми рисками. – 2018. – N 55(3). – С. 174 – 190.

292. Regulatory Consistency Assessment Programme (RCAP) — Analysis of Risk-Weighted Assets for Credit Risk in the Banking Book// https://www.bis.org/bcbs/publ/d363.pdf.

293. Basel III: Finalising Post-Crisis Reforms// https://www.bis.org/bcbs/publ/d424.pdf.

294. Rubchinsky A. Divisive-Agglomerative Classification Algorithm Based on the Minimax Modification of Frequency Approach// Working paper WP7/2010/07. – M.: NRU HSE. – 2010. – 48 p.

295. Rubchinsky A. Divisive-Agglomerative Algorithm and Complexity of Automatic Classification Problems// Working paper WP7/2015/09. – M.: NRU HSE. – 2015. – 44 p.

296. Rubchinsky A. Family of Graph Decompositions and Its Applications to Data Analysis// Working paper WP7/2016/09. – M.: NRU HSE. – 2016. – 60 p.

297. Rubchinsky A. A New Approach to Network Decomposition Problems// In: Kalyagin V., Nikolaev A., Pardalos P., Prokopyev O. (eds) Models, Algorithms, and Technologies for Network Analysis. NET 2016. Springer Proceedings in Mathematics & Statistics. – 197. Springer, Cham. – 2017.

298. Rubchinsky A. Graph Dichotomy Algorithm and Its Applications to Analysis of Stocks Market// In: Kalyagin V., Pardalos P., Prokopyev O., Utkina I. (eds) Computational Aspects and Applications in Large-Scale Networks. NET 2017. Springer Proceedings in Mathematics & Statistics. – 247. Springer, Cham. – 2017.

299. Rubchinsky A. Notion of Graph Entanglement and its Application for Stock Market Analysis// In: Proc. of 21st IEEE Conference on Business Informatics, Moscow, Russia, 15-17 July 2019. – P. 41 – 46.

300. Четверушкин Б.Н. Кинетические схемы и квазигазодинамическая система уравнений. – М.: МАКС Пресс. – 2004.

301. Елизарова Т.Г. Квазигазодинамические уравнения и методы расчета вязких течений. – М.: Научный мир. – 2007.

302. Zlotnik A., Lomonosov T. Verification of an Entropy Dissipative QGD-Scheme for the 1D Gas Dynamics Equations// Mathematical Modelling and Analysis. – 2019. – N 24(2). – P. 179 - 194.

303. Zlotnik A.A., Lomonosov T.A. Conditions for L2-Dissipativity of Linearized Explicit Difference Schemes with Regularization for 1D Barotropic Gas Dynamics Equations// Computational Mathematics and Mathematical Physics. – 2019. – N 59(3). – P. 452 - 464.

304. Злотник А.А., Ломоносов Т.А. Условия диссипативности линеаризованных явных разностных схем с регуляризацией для уравнений 1D баротропной газовой динамики// ЖВМиМФ. – 2019. – N 59(3). – С. 128 - 140.

305. Gentile C. Distributed Sensor Location through Linear Programming with Triangle Inequality Constraints// IEEE Transactions on Wireless Communications. – 2007. – N 6(7). – P. 4020 - 4027.

306. Salari M., Kattan L., Lam W.H.K., Lo H.P., Esfeh M.A. Optimization of traffic sensor location for complete link flow observability in traffic network considering sensor failure// Transportation Research Part B. – 2019. – N 121. – P. 216 – 251.

307. Leontief W.W. The Structure of American Economy// Harvard University Press. – 1941. – P. 1919 - 1929.

308. Page L., Brin S., Motwani R., Winograd T. The PageRank Citation Ranking: Bringing Order to the Web. – 1998.

309. Демин С.С. Проект системы раннего реагирования для устранения последствий разливов нефти в Северном Ледовитом океане// В кн.: Управление развитием крупномасштабных систем MLSD'2019 / Под общ. ред.: С.Н. Васильев, А. Цвиркун. – ИПУ РАН, 2019. – С. 977 - 979.

310. Smith A. An Inquiry into the Nature and Causes of the Wealth of Nations: A Selected Edition Adam Smith (Author). – Oxford: Kathryn Sutherland (Editor), Oxford Paperbacks. – 2008.

311. Rodriguez-Clare A. The division of labor and economic development// Journal of Development Economics. – 1996. – N 49. – P. 3 – 32.

312. Hummon N.P. A mathematical theory of differentiation in organizations// American Sociological Review. – 1971. – N 36(2). – P. 297 – 303.

313. Gavrilets S., Auerbach J., Vugt M.V. Convergence to consensus in heterogeneous groups and the emergence of informal leadership// Scientific reports. – 2016. – N 6. – P. 29704.

314. Perry L., Shrestha M.D., Vose M.D., Gavrilets S. Collective action problem in heterogeneous groups with punishment and foresight// Journal of Statistical Physics. – 2018. – N 172(1). – P. 293 – 312.

315. Smith J.M., Szathmary E. The Major Transitions in Evolution. – San Francisco: Freeman. – 1995.

316. Bratus A.S., Drozhzhin S., Yakushkina T. On the evolution of hypercycles// Mathematical biosciences. – 2018. – N 306. – P. 119 – 125.

317. Michod R.E., Viossat Y., Solari C.A., Hurand M., Nedelcu A.M. Life-history evolution and the origin of multicellularity// J. Theor. Biol. – 2006. – N 239. – P.257 – 272.

318. Ispolatov I., Madhok V., Doebeli M. Individual-based models for adaptive diversification in high-dimensional phenotype spaces// J. Theor. Biol. – 2016. – N 390. – P. 97 - 105.

319. Gavrilets S. Rapid Transition towards the Division of Labor via Evolution of Developmental Plasticity// PLoS Comput. Biol. – 2010. – N 6.

320. Rueffler C., Hermisson J., Wagner G.P. Evolution of functional specialization and division of labor// PNAS. – 2012. – N 109. – P. 326 – 335.

ПРИЛОЖЕНИЕ А

Таблица к подразделу 5.2

Таблица А.1 – Сцепленности фрагмента фондового рынка S&P-500

1-й квартал 1991 г.

1371	1379	355	21	99	128	48	227	145	79	95	79	29	58	88
66	1350	17	147	15	15	26	25	13	13	239	17	34	14	20
8	17	80	84	206	57	291	88	42	54	388	702	9	956	124
30	11	6	10	1229	1555	705	506	274	37	995	42	252	113	1326
723	1616	114	423	20	26	461	386	402	52	247	685	471	329	1246
1217	1421	51	445	120	21	16	62	38	42	552	570	504	1256	49

2-й квартал 1991 г.

669	102	88	417	94	171	352	630	593	338	135	761	1659	1258	377
552	29	27	501	1508	865	247	61	268	472	213	410	421	148	34
217	107	585	251	271	347	210	68	413	1008	1315	622	1306	891	1127
7	37	16	304	20	48	274	448	153	1756	1805	1393	709	1460	47
1121	526	575	22	35	56	6	24	35	7	7	105	172	85	511
855	666	106	879	880	270	656	265	472	500	1562	1366	388	99	35	218

3-й квартал 1991 г.

49	172	88	480	1120	1120	1131	453	154	158	252	290	247	336	1275
453	91	43	109	207	222	90	106	237	250	500	38	82	87	39
76	812	819	1424	525	892	52	35	55	27	155	102	259	140	810
205	359	16	70	57	21	383	412	244	892	1595	602	1136	544	209
15	34	116	5	12	78	370	499	414	251	22	140	79	26	86
64	42	69	27	682	555	569	811	1351	1733	81	647	811	1255	922	108	67

4-й квартал 1991 г.

367	1399	554	265	687	1071	63	220	210	946	1294	297	92	1430	331
252	673	38	236	103	27	171	42	25	57	186	561	447	1406	900
94	94	416	486	129	10	209	230	182	61	665	11	383	122	36
34	140	66	817	84	58	258	31	102	76	107	26	5	10	20
30	44	597	1064	366	351	367	818	1157	671	664	445	178	417	388
1057	759	641	218	413	52	48	48	109	676	1135	14	54	161	15	32	6

1-й квартал 1992 г.

1797	1801	1823	1691	1789	1573	608	458	556	222	201	142	129	42	258
471	74	56	11	70	16	126	315	788	65	668	576	590	266	62
40	524	255	587	142	35	738	100	282	594	189	1436	44	368	64
18	180	118	42	651	1635	835	816	506	42	52	243	406	270	856
1351	990	912	1464	1787	1460	1121	1585	31	413	365	32	280	511	828
25	1658	331	573	308	33	13	57	206	743	26	160	246	101	288	62

2-й квартал 1992 г.

123	156	114	61	15	313	729	774	302	1335	181	76	15	1651	177
533	61	28	73	16	35	421	116	351	80	42	87	214	413	61
177	547	621	1471	906	859	1022	96	367	270	305	644	127	596	66
343	717	385	14	148	215	282	409	58	12	102	58	355	85	189
473	92	74	281	1222	1418	1454	688	1110	1155	138	1599	200	392	17
619	230	16	94	481	357	82	322	28	88	353	517	51	102	21	154

Продолжение таблицы А.1

3-й квартал 1992 г.

376 98 121 12 12 15 19 1068 503 682 1195 996 1398 1090 937
197 37 585 1226 312 410 698 1150 538 70 254 1618 1537 253 959
813 157 196 45 59 16 8 7 30 106 466 282 326 1217 266
398 72 382 60 1519 59 44 263 11 631 1644 159 1263 147 111
439 252 980 1502 263 294 127 133 10 133 81 133 73 129 139
120 861 889 17 69 661 523 161 13 195 384 1032 1711 1512 148 741 306

4-й квартал 1992 г.

205 114 39 210 568 856 20 496 92 42 52 207 336 76 36
76 235 36 519 712 1403 69 919 238 95 245 573 1000 1286 408
18 308 833 198 43 149 27 45 971 249 531 565 86 700 213
54 33 209 159 1047 455 202 269 763 48 289 318 162 507 20
106 161 154 161 34 119 135 1259 36 125 361 481 595 1156 770
301 278 153 810 204 98 74 121 58 761 23 33 94 339 214 107 52

1-й квартал 1993 г.

276 302 118 23 506 7 649 223 390 124 199 127 125 158 59
59 228 181 86 989 65 462 212 15 601 1361 1109 754 652 1007
374 938 1505 22 218 101 125 33 36 123 24 45 170 49 76
374 584 789 81 24 64 54 28 264 45 213 309 67 117 722
1119 98 39 24 46 734 323 214 618 1262 206 14 465 53 1550
1310 380 1634 1729 354 1004 36 413 268 97 509 949 450 114 1315

2-й квартал 1993 г.

1226 349 362 1094 289 146 750 75 713 538 1254 176 58 153 88
146 153 247 583 595 736 144 167 401 974 60 113 18 157 112
392 355 629 1142 82 468 1346 1291 27 413 1539 710 1535 991 165
1722 177 226 43 116 289 54 298 342 108 1032 121 678 106 138
52 635 17 624 1019 473 713 133 105 345 77 145 49 117 40
300 859 1094 629 583 378 1462 687 653 319 1313 45 65 163 760 388

3-й квартал 1993 г.

443 28 150 207 162 1245 428 166 708 161 548 1370 695 649 894
1415 1637 480 279 761 1370 369 63 23 20 120 618 831 166 162
101 214 113 332 346 124 567 105 186 251 155 510 61 141 541
945 717 278 1115 1339 663 136 152 172 362 172 243 117 253 648
897 361 70 428 354 654 184 209 61 472 402 216 623 65 184
1440 159 41 191 1216 623 1608 774 552 448 338 309 108 105 315 556 120

4-й квартал 1993 г.

33 440 413 1079 753 824 630 1226 1379 623 1626 944 1384 1557 81
100 86 1351 85 138 331 284 125 368 44 640 982 1303 1181 247
415 859 1043 120 76 321 462 19 179 81 51 603 106 167 431
49 999 169 27 169 126 107 893 493 204 1415 773 1524 1491 375
536 1262 568 587 411 1370 198 79 860 335 351 223 856 193 103
1572 717 102 77 403 1265 95 118 41 254 375 1260 35 552 57 695 1241

1-й квартал 1994 г.

924 35 1099 544 1400 293 38 131 109 444 220 198 328 389 1482
125 210 124 421 334 14 144 7 87 43 125 300 1596 143 51
320 1176 432 1586 355 118 1641 1557 375 848 171 815 49 41 181
383 565 173 171 10 5 23 463 94 861 695 42 41 179 362
161 189 440 169 597 117 201 1094 30 114 417 227 1494 109 138
63 1491 976 1197 1354 214 1034 116 959 182 376 425 1054 135 248

Продолжение таблицы А.1

2-й квартал 1994 г.

19 20 70 21 177 1812 69 31 20 98 230 12 168 37 35
69 362 378 49 82 32 341 41 123 159 138 52 700 646 215
1627 1607 696 1580 990 846 74 834 364 563 31 275 45 420 187
202 258 219 314 319 5 133 1161 243 130 479 346 367 13 53
358 130 1018 87 132 184 264 972 834 859 102 66 1299 333 646
686 293 763 393 620 1350 371 224 437 1551 157 104 8 351 46 11

3-й квартал 1994 г.

137 19 53 574 56 407 416 1566 1008 879 814 264 521 167 756
162 120 889 296 294 1231 391 630 99 1201 612 1330 1177 1783 893
1307 169 238 460 376 264 1023 1628 208 220 880 17 61 352 423
386 580 276 925 725 42 337 18 178 337 276 226 87 969 1473
856 70 228 540 57 354 304 291 1043 967 91 613 1267 729 1238
722 98 1060 289 296 722 234 1306 229 29 42 21 38 29 33 81 923

4-й квартал 1994 г.

487 71 62 196 803 182 114 38 185 50 540 382 140 362 149
514 98 36 104 98 17 23 901 42 101 37 182 58 65 61
49 943 1011 429 316 125 728 950 1197 507 1229 862 123 214 418
176 1590 267 856 276 100 234 245 145 32 95 5 6 5 17
10 7 17 106 755 168 378 343 591 156 620 797 206 436 160
540 686 10 68 326 286 225 127 159 56 509 417 12 221 64 574 599

1-й квартал 1995 г.

28 708 713 1560 846 1470 1748 1398 990 1223 191 370 464 345 335
1071 279 103 102 406 497 771 232 200 32 224 452 188 371 178
271 846 298 513 505 782 47 150 438 48 169 57 1596 1682 772
319 78 929 301 158 81 71 873 739 337 454 192 80 876 1562
1710 506 62 54 1278 143 1077 1011 868 693 302 992 404 893 67
399 684 588 434 38 96 50 109 62 1309 862 281 675 113 1366

2-й квартал 1995 г.

1344 907 678 759 194 204 324 169 653 600 13 468 522 793 664
1333 611 1339 514 1254 1247 101 266 193 1194 203 334 534 573 1015
61 281 844 367 97 39 100 109 640 118 34 65 349 153 909
131 971 958 224 970 286 461 1537 1023 244 163 1131 142 279 612
881 1492 1328 1750 505 518 126 40 520 1680 578 477 164 300 34
179 126 34 660 510 700 65 47 9 17 160 82 38 746 586 605

3-й квартал 1995 г.

368 723 1293 581 1487 518 1589 94 34 32 34 91 11 19 83
176 189 231 1762 748 333 309 308 139 113 368 1186 531 54 199
30 417 141 970 1033 562 71 33 1313 257 678 144 408 551 877
762 1022 493 816 227 118 181 142 1089 124 999 1273 1487 1178 1486
193 846 119 234 550 567 169 1729 570 384 339 159 447 82 530
31 746 475 556 164 26 65 344 41 494 58 38 78 78 186 187 34

4-й квартал 1995 г.

294 120 204 40 197 16 47 17 33 63 8 610 410 72 122
84 695 89 871 1325 637 79 1207 350 110 101 46 224 307 293
99 415 786 20 22 691 656 1615 61 452 241 105 197 393 517
1390 91 265 23 312 323 128 249 50 228 151 684 16 1051 437
82 16 12 172 535 58 425 239 93 354 48 67 81 498 244
220 71 220 32 72 124 104 107 165 67 53 132 314 1049 66 151 270

Продолжение таблицы А.1

1-й квартал 1996 г.

360 628 469 394 1436 1173 963 74 372 51 31 42 111 10 100
8 7 267 50 133 107 188 561 210 79 550 271 678 158 1067
285 276 921 1305 513 24 42 12 6 10 74 21 59 80 575
320 151 901 125 264 79 94 531 486 125 116 528 940 581 1082
1216 978 35 170 1009 104 115 25 285 479 70 944 316 81 305
267 779 425 478 98 705 1327 363 736 871 387 765 528 52 119 85

2-й квартал 1996 г.

428 349 465 419 111 29 68 47 320 751 412 225 60 57 93
150 50 65 159 18 81 197 34 176 400 1517 1218 841 1386 1288
95 252 20 399 89 616 45 29 273 240 698 863 534 1152 1343
296 136 324 125 17 57 336 176 17 292 130 44 209 22 164
202 1225 259 105 314 310 897 299 14 15 12 504 198 457 421
19 18 95 369 129 739 938 298 370 1092 780 638 337 711 393 285

3-й квартал 1996 г.

192 636 1064 351 61 149 119 644 25 3 5 4 93 257 29
13 16 4 408 111 27 105 797 135 72 15 165 79 1216 148
360 168 1561 353 222 74 32 9 11 14 10 38 507 269 23
1113 55 295 566 300 326 110 191 264 231 385 855 171 53 742
20 459 1365 58 19 43 36 21 107 168 472 321 285 171 8
62 71 87 144 23 52 34 16 26 83 33 665 584 160 99 318 64

4-й квартал 1996 г.

930 144 340 295 14 45 50 291 70 389 134 562 87 295 311
99 178 1136 375 42 19 322 9 73 594 465 187 224 688 14
293 182 368 132 334 226 798 356 176 231 152 104 48 78 109
18 32 26 305 336 92 9 16 21 139 152 80 40 412 106
70 233 32 45 70 496 493 1016 61 133 405 156 74 228 44
24 147 34 12 16 61 41 417 53 80 35 240 10 17 27 28 22

1-й квартал 1997 г.

111 916 140 358 40 82 432 323 1119 167 152 46 93 903 31
19 917 134 125 165 1447 174 457 806 356 1089 43 30 12 20
462 455 69 157 1029 1206 132 173 70 137 144 196 72 180 166
29 18 50 122 540 426 111 214 490 500 219 1567 1177 92 246
195 697 1070 141 878 695 96 30 167 132 277 1434 717 998 655
446 685 547 926 93 119 133 25 22 265 44 405 530 483 66

2-й квартал 1997 г.

103 87 8 9 322 7 14 151 10 16 1119 1086 1235 77 755
1683 1776 629 46 85 338 1138 178 161 50 63 36 401 994 85
844 140 80 39 10 13 4 5 9 19 10 9 6 7 59
34 1393 538 752 162 54 73 272 195 427 16 451 44 59 1187
105 324 1383 57 363 57 298 9 32 1399 167 586 75 31 51
54 54 272 24 133 32 47 57 163 191 456 174 99 125 52 125

3-й квартал 1997 г.

170 405 114 541 893 722 286 1073 91 163 564 98 17 336 814
35 115 373 12 82 9 322 390 428 33 448 316 136 801 326
427 341 48 653 634 408 236 7 439 801 204 678 1249 113 728
115 19 10 63 18 287 255 290 681 71 683 62 1596 259 822
148 630 739 276 53 129 188 547 721 74 28 131 93 351 60
50 225 479 775 659 128 294 68 14 236 697 98 41 458 113 1437 933

Продолжение таблицы А.1

4-й квартал 1997 г.

1736 578 200 486 970 1161 434 18 77 83 1365 79 8 41 1242
772 682 317 1399 10 348 330 499 21 124 752 383 459 1743 75
1578 71 290 605 529 16 21 1262 150 27 16 68 166 201 1559
423 695 224 124 60 1309 294 45 7 75 12 40 14 105 68
37 384 348 334 258 35 135 28 495 183 95 112 31 136 603
112 921 249 31 74 55 129 78 74 203 627 256 234 105 51 162 694

1-й квартал 1998 г.

115 662 188 835 1533 110 5 32 1131 107 82 9 43 51 23
49 41 39 33 80 193 556 179 277 223 167 207 306 84 11
72 165 200 529 774 449 28 106 108 18 79 42 80 7 239
70 26 40 774 582 296 699 584 779 478 84 912 18 471 273
940 1306 433 132 251 56 70 311 662 39 57 42 6 22 76
232 19 128 8 41 278 386 154 89 73 525 127 272 463 712

2-й квартал 1998 г.

381 1719 1039 676 925 229 550 502 255 69 92 101 198 119 578
613 74 180 62 82 42 286 500 31 427 575 290 227 55 229
41 34 194 251 104 89 1687 92 460 112 205 439 806 1186 130
363 220 25 83 18 192 872 18 55 501 43 120 349 313 234
446 485 11 71 58 80 57 68 179 74 1060 402 161 260 175
164 32 121 27 154 300 1002 969 147 1071 102 110 176 245 249 419

3-й квартал 1998 г.

47 587 135 905 1198 181 9 233 729 125 123 90 70 17 19
86 506 103 519 107 701 1198 1082 633 16 27 7 18 14 15
24 9 40 264 58 16 33 164 40 189 285 1047 996 52 208
1258 437 239 273 722 237 102 389 192 261 60 526 252 310 30
13 48 6 19 5 111 50 18 10 14 56 8 518 346 505
1392 1607 229 230 17 438 123 17 90 855 58 78 189 356 699 164 64

4-ый квартал 1998 г.

37 48 150 228 8 95 21 175 8 9 18 19 250 449 96
107 12 3 27 52 15 13 40 14 46 55 23 44 137 335
89 207 397 1309 218 164 881 115 1251 101 12 14 117 15 156
84 469 170 384 453 246 173 1202 1420 1019 167 86 486 139 779
682 158 819 847 120 163 66 787 30 725 75 675 113 297 395
946 137 59 578 518 578 104 1454 26 37 323 349 43 624 55 49 10

1-й квартал 1999 г.

610 129 195 963 36 187 35 157 33 353 36 280 84 91 49
316 824 48 429 51 136 56 273 13 99 60 81 55 120 96
173 23 55 516 99 906 249 85 38 51 256 33 34 130 70
64 1324 533 19 37 1224 418 24 1095 236 159 262 211 547 232
50 1281 949 1068 1762 128 1066 488 188 109 37 11 776 790 814
357 65 1097 179 29 157 90 419 1224 657 734 679 23 452 352

2-й квартал 1999 г.

17 314 53 512 921 222 49 165 54 491 785 354 38 12 1660
521 52 105 50 196 29 15 33 38 64 6 143 191 1172 713
889 1504 1370 1018 661 35 480 14 1121 107 44 534 94 1084 672
425 1693 190 52 849 497 144 801 18 15 9 75 23 7 194
22 346 351 738 8 331 1198 369 143 160 135 16 42 92 607
342 1048 735 150 174 1034 1273 1378 188 271 107 286 421 107 168 84

Продолжение таблицы А.1

3-й квартал 1999 г.

1192 1126 224 458 360 136 169 30 269 75 181 628 51 91 78
245 41 80 585 569 228 417 52 484 117 26 127 203 330 35
22 39 7 161 790 63 93 155 326 258 71 18 108 35 347
931 501 237 728 1048 185 1276 620 84 65 210 177 22 30 281
45 701 35 663 38 59 81 60 154 82 96 170 178 295 251
1192 19 444 204 262 42 27 58 145 409 18 1271 564 52 28 89 15

4-й квартал 1999 г.

54 72 314 56 13 1041 638 368 35 30 51 1276 110 383 341
550 327 60 12 84 1460 1451 632 34 28 13 354 55 769 131
421 47 113 115 242 92 223 946 109 692 1080 25 619 435 252
37 479 268 127 498 107 245 549 288 17 50 132 54 45 26
13 34 325 65 1060 336 1380 86 443 522 23 14 98 24 64
30 19 173 80 82 464 1216 812 1610 1264 425 676 1352 1319 1401 185 126

1-й квартал 2000 г.

173 135 94 222 205 796 1265 33 96 251 225 359 842 277 1589
1177 1220 1062 313 658 75 144 10 37 11 49 24 92 138 196
389 863 1164 15 169 574 967 556 10 27 265 174 24 47 72
8 26 9 102 57 22 507 88 516 385 141 90 148 95 5
52 36 22 407 211 400 103 843 185 38 323 374 115 42 200
61 819 338 167 289 96 48 36 72 29 70 302 1242 1600 1430 210

2-й квартал 2000 г.

305 538 622 114 21 80 100 175 135 1038 31 236 399 19 15
23 17 9 14 116 40 37 64 21 52 21 29 67 797 167
1198 616 161 163 548 18 148 29 44 46 1150 1245 317 253 183
42 80 763 40 38 14 3 27 7 20 3 3 471 91 55
934 304 137 514 348 388 179 118 613 16 186 436 1062 428 267
320 272 32 52 16 19 206 158 49 105 65 22 59 1301 66 415

3-й квартал 2000 г.

117 162 36 72 453 717 1469 68 154 692 264 228 128 27 64
373 42 74 637 1420 179 117 172 16 288 22 46 116 83 87
156 34 35 312 138 216 132 31 135 139 14 282 13 74 102
484 560 1071 249 191 341 167 140 132 76 100 41 138 434 71
48 1023 1416 67 51 354 352 1264 1143 1088 100 136 49 59 151
236 37 65 318 135 40 234 520 5 20 95 116 14 17 86 20 19

4-й квартал 2000 г.

56 224 155 121 52 243 167 503 198 388 509 29 859 41 472
1282 996 57 42 249 106 202 616 238 36 999 194 829 12 219
71 298 94 351 145 247 324 556 328 83 99 60 1732 57 113
221 135 603 193 614 35 672 328 160 10 237 16 197 515 1088
53 1568 1511 809 92 364 308 476 1034 241 594 53 383 291 34
11 79 106 43 55 113 30 40 47 189 160 99 191 17 59 10 40

1-й квартал 2001 г.

52 91 936 20 42 176 157 125 44 98 39 13 9 16 19
242 317 683 914 1341 1007 186 1602 1739 1340 1240 555 494 463 345
198 416 743 953 122 111 384 87 402 250 286 165 61 33 132
349 782 163 319 1426 1474 1725 429 1265 99 37 10 718 754 40
144 31 78 36 1255 742 85 767 44 374 20 441 241 219 540
43 103 20 22 5 271 278 44 12 13 62 53 131 487 26

Продолжение таблицы А.1

2-й квартал 2001 г.

267 485 629 185 19 84 104 181 150 1124 32 222 417 19 20
30 17 10 16 103 35 42 60 20 49 19 21 131 839 166
1281 663 153 144 533 18 148 30 44 37 1086 1173 367 262 157
42 72 758 55 56 9 4 31 5 16 3 4 396 76 53
992 232 154 479 335 450 241 118 685 17 180 431 891 514 284
358 244 29 55 16 20 202 122 51 111 63 20 66 1323 77 418

3-й квартал 2001 г.

116 66 48 158 212 32 240 1076 1647 1075 64 373 21 43 242
100 81 279 79 151 148 28 87 15 420 475 119 118 118 302
601 92 459 104 53 251 210 298 504 597 534 169 375 30 248
48 33 405 221 28 17 60 12 78 40 85 1063 156 418 24
133 324 21 290 285 47 405 55 67 56 25 87 9/11 terrorist attack
56 14 129 8 23 4 20 10 9 7 14 135 149 184

4-й квартал 2001 г.

96 334 251 363 185 120 458 26 65 150 85 53 27 167 233
498 6 193 21 1040 1490 343 1361 1033 276 417 1272 235 627 94
464 287 352 43 909 1579 215 42 7 160 97 64 131 109 15
3 3 6 345 56 10 186 251 27 291 56 241 111 727 1008
1280 621 1553 1541 464 36 230 20 19 69 71 155 6 338 540
1386 920 42 1098 123 40 48 18 277 921 550 61 112 49 46 44 31

1-й квартал 2002 г.

530 184 1276 185 215 345 27 14 31 492 35 98 83 14 302
429 788 181 23 210 41 55 69 38 367 632 278 1270 867 32
437 49 133 339 60 20 574 426 301 243 302 947 167 1681 430
847 402 110 245 490 110 13 11 420 95 106 696 598 56 285
872 86 350 13 8 10 5 19 11 19 80 220 58 459 812
541 1281 669 22 37 109 282 644 415 412 234 343 27 597 471

2-й квартал 2002 г.

1408 1703 1556 167 574 47 14 70 398 15 39 39 20 273 10
26 491 81 1594 1740 1102 745 448 1176 1049 33 15 1021 152 140
122 103 157 114 128 154 211 482 755 568 191 36 162 1143 381
547 34 66 265 120 417 84 931 377 379 433 279 84 288 23
219 73 223 191 23 185 7 10 14 5 40 73 154 77 14
7 12 17 122 276 1525 1037 16 11 539 43 722 64 130 880 126

3-й квартал 2002 г.

12 18 58 11 30 17 10 124 1141 1576 43 8 25 31 128
86 18 144 22 9 5 5 33 8 6 51 443 32 65 154
409 11 49 215 41 82 691 377 32 432 592 879 1391 42 498
55 14 41 112 171 816 639 266 1491 381 1051 32 15 194 516
245 411 72 51 11 12 83 20 6 135 111 1062 401 1472 409
759 321 23 53 646 65 15 107 676 34 56 30 4 5 3 111 27

4-й квартал 2002 г.

67 71 1093 344 465 24 287 256 20 77 56 1109 654 1211 409
62 230 43 167 148 74 100 4 10 90 19 8 90 71 119
121 559 181 52 1292 84 194 991 634 490 585 485 76 1379 1390
818 1717 1790 1488 1777 1730 756 93 63 8 61 20 19 83 26
9 40 24 65 237 923 415 29 94 70 25 111 17 127 13
94 57 144 10 228 62 108 108 407 31 91 1000 95 10 14 442 1684

Продолжение таблицы А.1

1-й квартал 2003 г.

71 30 1370 89 27 38 12 6 3 71 33 12 7 90 46
23 304 43 636 41 301 30 48 17 404 75 9 10 13 12
43 7 11 8 54 576 767 109 341 262 57 80 19 56 20
84 77 46 47 1022 505 891 221 498 318 59 67 368 86 36
3 80 1258 151 549 135 26 36 40 256 16 722 10 55 425
61 70 14 441 13 199 604 4 13 4 15 5 5 23 97

2-й квартал 2003 г.

448 184 156 45 46 102 157 412 115 47 428 884 1142 926 75
1509 71 36 714 23 58 518 168 49 9 208 91 32 1183 59
113 25 74 27 168 33 1387 121 106 873 567 13 86 153 625
119 452 59 275 82 629 105 42 349 55 239 280 76 201 351
1294 123 917 1084 8 19 4 140 45 13 20 27 108 92 112
129 508 364 1413 55 25 934 834 506 518 22 1233 670 90 233 48

3-й квартал 2003 г.

71 94 940 323 1148 1108 991 43 251 59 53 25 15 13 7
98 160 370 1512 1593 492 725 828 234 956 549 185 562 1208 466
763 446 113 106 204 712 61 57 476 67 478 526 353 590 691
156 1451 631 571 102 19 86 19 131 69 46 90 160 383 333
461 59 35 405 381 20 16 76 45 911 583 192 117 463 1010
309 190 55 359 1249 721 237 1496 160 121 9 77 44 139 224 682 120

4-й квартал 2003 г.

254 165 30 143 277 127 298 8 47 72 23 25 7 11 32
8 76 9 471 315 215 1478 387 387 422 498 188 760 30 696
131 136 76 311 511 116 240 69 251 295 161 54 713 116 121
1572 1623 288 34 669 59 320 541 59 502 50 38 253 543 104
10 45 269 995 4 15 103 74 143 203 25 379 42 138 128
122 298 644 34 1393 135 99 180 935 28 41 154 15 11 172 11 11

1-й квартал 2004 г.

1660 41 30 1047 1536 96 72 303 59 50 20 224 14 123 104
1462 22 262 501 444 82 361 1085 53 28 44 348 892 683 310
498 412 1199 11 25 23 19 71 65 1271 491 1603 914 1325 304
412 153 73 21 1074 465 284 98 472 194 119 71 178 179 89
1207 29 101 26 37 100 89 25 33 194 97 498 4 5 4
5 6 5 24 38 65 183 179 599 113 84 81 167 296 105 68

2-й квартал 2004 г.

115 375 97 40 104 16 71 9 26 45 24 175 56 284 1481
9 21 91 194 22 11 131 241 333 94 147 151 1700 254 420
10 9 22 14 117 54 46 11 37 306 681 22 12 71 23
423 13 411 405 380 85 856 97 44 488 3 10 27 19 43
31 70 31 14 822 1097 483 400 1226 1574 210 185 125 428 48
874 651 478 83 30 40 60 109 42 183 311 150 115 129 768 1130

3-й квартал 2004 г.

240 1011 139 15 169 552 388 29 82 85 125 143 203 244 393
56 35 141 42 9 57 479 132 235 808 137 397 49 147 69
174 5 42 71 786 260 506 71 381 46 46 10 5 9 11
50 8 208 237 127 34 390 29 28 13 7 9 26 568 1109
418 163 310 892 27 87 37 15 20 12 15 786 529 64 40
257 46 371 412 83 139 174 62 552 690 304 314 555 161 453 34 227

Продолжение таблицы А.1

4-й квартал 2004 г.

121 127 122 7 15 23 38 26 8 19 68 295 44 252 456
72 607 84 39 116 278 145 153 66 31 94 731 821 756 150
35 13 158 8 8 5 7 14 52 61 27 144 245 80 77
549 476 824 577 267 63 76 613 41 1348 1612 274 69 296 100
483 32 264 57 17 279 18 162 115 241 146 303 172 75 110
848 42 28 216 1209 936 269 107 100 17 554 377 143 76 380 1315 172

1-й квартал 2005 г.

20 265 401 721 777 1130 74 95 9 10 12 9 38 10 31
90 79 873 562 72 347 307 25 43 470 557 106 58 196 26
156 245 57 73 83 107 126 25 3 39 12 11 8 33 406
167 346 753 450 37 53 213 173 127 302 118 98 77 35 294
201 1007 167 209 30 125 90 1246 833 730 443 73 162 1172 289
61 357 37 53 85 303 27 80 135 624 662 319 20 210 564

2-й квартал 2005 г.

1515 1233 557 1016 155 395 119 249 279 1184 511 11 692 489 10
34 7 17 3 6 13 51 289 302 50 131 15 195 464 1052
245 66 986 234 54 147 814 115 15 136 1252 726 105 189 441
1085 6 255 128 260 28 66 28 5 3 8 13 24 58 39
448 201 93 31 238 367 503 490 144 7 19 92 45 52 275
43 79 27 36 363 485 70 39 546 466 45 401 181 83 179 568

3-й квартал 2005 г.

580 417 375 28 194 14 98 60 758 16 339 265 430 286 50
59 12 7 29 101 18 1423 166 72 88 34 247 19 268 32
104 305 1168 619 1184 276 166 69 512 607 191 93 190 17 5
48 83 75 368 36 509 409 270 279 63 184 51 56 196 149
808 128 361 122 104 84 790 187 224 7 100 253 130 27 437
274 491 154 92 257 14 34 56 26 14 33 24 28 16 63 537 445

4-й квартал 2005 г.

153 164 396 48 463 1315 418 57 125 806 19 15 6 35 52
94 89 185 428 151 67 348 413 98 132 244 166 594 143 131
289 49 630 35 21 48 10 29 30 153 430 808 181 120 260
244 208 253 430 137 500 552 181 15 8 51 68 18 15 42
40 1268 158 49 86 163 389 61 135 49 19 133 399 710 308
853 453 108 190 1485 83 309 1188 130 151 541 246 366 36 37 62 49

1-й квартал 2006 г.

10 360 70 386 54 326 336 326 32 17 202 358 511 39 95
65 723 488 689 915 11 9 303 92 9 535 361 108 133 22
551 1159 1386 1140 61 304 41 1352 322 1531 65 284 238 165 1592
77 681 18 43 15 10 76 24 24 432 117 70 55 16 598
1757 415 1462 1636 1103 36 2 5 5 5 4 5 3 2 2
4 9 14 2 11 8 5 57 71 47 447 1098 605 388 250

2-й квартал 2006 г.

19 81 371 721 1470 420 44 140 11 117 184 369 638 71 287
547 1149 859 59 75 200 31 72 33 507 122 299 155 200 122
142 56 257 831 451 394 619 70 142 194 112 1743 744 135 1228
39 5 39 8 18 93 11 14 9 93 54 122 319 147 836
956 1599 50 34 94 247 110 97 52 241 53 39 16 41 35
115 459 451 368 324 695 214 251 23 775 1004 29 654 466 22 260

Продолжение таблицы А.1

3-й квартал 2006 г.

16 56 134 461 101 19 55 61 664 464 193 60 14 254 79
4 36 22 237 274 45 24 21 331 58 542 875 69 35 54
36 92 16 77 175 178 604 610 404 431 33 674 93 1233 585
699 1052 33 12 32 70 281 29 16 17 551 218 399 89 1071
975 253 1183 432 646 72 22 65 100 53 120 27 189 458 813
350 33 11 135 246 14 5 23 21 312 127 1287 466 1121 250 226 10

4-й квартал 2006 г.

452 414 283 272 156 298 12 90 910 69 395 302 69 62 18
914 467 26 44 16 306 322 19 94 208 409 796 141 482 1482
379 328 643 71 215 309 20 670 896 306 446 481 793 501 206
88 537 976 364 181 749 37 82 143 98 105 138 291 1550 1610
211 151 12 85 22 260 456 286 545 198 15 33 80 236 43
742 454 493 404 35 177 80 435 238 72 339 114 37 151 14 228 259

1-ый квартал 2007 г.

582 549 980 1279 1046 1343 136 174 92 361 138 236 209 319 276
79 46 23 30 18 39 442 82 25 125 163 26 53 69 137
206 149 173 14 21 76 15 30 26 7 127 63 428 205 128
1602 625 1348 1340 39 23 134 594 44 66 51 354 41 39 159
27 108 289 2 10 7 3 15 25 14 4 11 34 295 708
112 79 1249 162 1112 19 432 579 1375 27 6 10 115 197 66

2-й квартал 2007 г.

104 367 338 679 577 196 42 138 9 525 457 256 109 69 356
25 295 948 30 32 214 48 275 2 157 59 278 117 12 356
1138 153 1738 216 358 754 144 59 128 731 13 1054 782 348 685
296 278 1322 748 508 43 286 122 94 95 380 375 70 56 527
369 160 71 20 124 336 880 292 474 101 48 74 119 27 24
60 57 27 11 150 707 24 7 275 735 653 59 492 19 605 967

3-й квартал 2007 г.

788 949 678 1582 1595 1097 48 793 201 1358 74 227 300 516 856
972 214 1079 128 267 18 454 134 252 45 12 88 37 49 8
27 69 6 6 4 13 119 159 124 46 192 116 62 155 1505
342 411 290 208 585 26 179 769 82 239 542 71 218 120 1462
91 507 1275 162 165 161 1176 25 11 281 198 822 518 416 34
101 30 19 94 215 140 65 78 11 21 113 10 131 14 76 345 111

4-й квартал 2007 г.

44 746 1610 224 37 143 15 23 196 80 243 127 242 94 1368
29 446 166 1083 264 25 76 23 30 294 164 82 12 68 312
160 57 76 212 269 20 97 845 42 83 8 3 19 1252 783
266 36 43 28 518 899 23 120 1147 283 105 740 234 155 75
71 90 465 400 16 11 18 201 366 10 3 51 39 283 714
350 177 33 118 38 17 15 18 133 417 359 119 255 271 201 444 168

1-й квартал 2008 г.

750 218 73 78 15 11 40 8 2 3 11 10 8 9 60
43 161 35 88 48 13 31 47 9 16 10 14 50 201 13
146 20 19 46 128 1326 1381 1312 83 9 23 87 223 204 11
73 54 44 82 873 30 234 180 547 1149 1795 366 36 1210 245
96 90 300 322 191 1060 121 10 6 10 10 231 99 282 28
183 202 4 4 947 5 4 4 29 28 10 31 327 63 214 67

Продолжение таблицы А.1

2-й квартал 2008 г.

1481 8 33 240 66 79 594 218 120 162 699 318 643 191 24
72 73 309 64 146 994 9 164 401 595 45 35 25 132 274
191 581 43 30 204 353 66 259 533 7 103 578 217 120 114
566 60 37 42 11 31 333 1189 62 68 15 58 68 17 225
19 6 17 17 135 404 295 40 136 71 5 30 10 23 156
74 11 6 11 71 216 278 92 72 19 61 219 41 9 18 17

3-й квартал 2008 г.

60 9 9 6 200 298 6 16 56 56 111 182 214 131 202
527 794 192 371 150 429 1318 25 15 26 33 63 32 24 79
1060 1250 1722 1220 1703 1250 1322 1254 273 127 130 30 17 16 13
10 5 7 12 6 7 5 9 4 7 15 123 176 178 1020
16 962 256 101 35 144 402 27 45 715 117 26 172 264 166
22 12 20 54 87 1511 720 423 213 914 1264 254 1028 173 26 1513 1783

4-й квартал 2008 г.

1853 1874 129 242 61 99 41 26 49 415 190 25 6 3 1578
16 6 1605 1025 211 49 1366 17 2 62 382 173 26 14 13
99 175 67 11 10 5 9 15 17 97 199 1110 463 9 5
90 1395 303 229 21 30 84 89 20 18 1323 668 1422 710 4
18 7 10 65 1476 404 16 27 38 47 52 227 216 69 253
405 518 361 436 151 91 1494 1329 1772 93 143 288 154 305 593 573 1793

1-й квартал 2009 г.

196 694 81 73 678 184 1330 228 24 518 109 1208 503 11 4
11 8 11 7 11 6 25 5 180 228 690 56 334 465 1014
882 481 387 1387 1562 163 1684 876 747 428 133 109 535 6 40
373 208 313 753 31 487 118 9 11 5 7 4 508 107 135
41 59 20 6 36 10 31 11 398 613 529 25 112 322 1070
1043 1664 715 10 11 29 1247 1372 82 79 8 286 435 35 658

2-й квартал 2009 г.

559 75 215 392 1345 8 14 23 29 16 64 20 9 31 21
17 75 11 22 138 14 15 147 574 117 104 148 578 1241 205
1119 684 1203 339 58 201 139 190 46 126 49 141 706 262 714
115 418 78 199 6 72 43 754 195 76 32 279 878 1052 85
41 68 103 288 195 211 8 54 59 35 71 86 211 143 172
1701 719 141 317 266 108 183 20 15 6 41 6 28 40 43 30

3-й квартал 2009 г.

17 383 1368 320 1298 617 866 96 11 21 22 3 56 3 115
258 9 51 55 25 29 167 85 839 517 212 21 23 54 481
35 366 265 113 453 296 202 445 94 25 781 249 71 21 144
680 154 399 1003 488 1049 1535 40 473 350 242 16 60 463 10
11 19 73 982 407 336 95 989 297 168 1236 15 11 155 11
8 210 11 26 10 9 7 6 11 3 10 168 33 70 1278 84 512

4-й квартал 2009 г.

98 49 321 68 61 11 137 389 317 31 19 274 35 37 10
5 5 89 14 64 22 139 96 75 432 945 621 124 94 81
55 664 333 37 18 18 68 87 174 28 29 65 26 7 8
5 9 5 8 28 24 60 922 26 161 487 298 488 26 17
26 12 611 72 1171 389 36 140 512 363 286 10 189 145 51
58 459 1179 760 79 813 828 147 122 128 52 10 32 6 185 36 7

Продолжение таблицы А.1

1-й квартал 2010 г.

13 239 350 64 950 1346 167 43 49 269 538 79 29 16 152
89 133 33 151 106 1393 558 233 11 5 10 33 20 4 9
11 65 175 5 839 716 80 101 188 1159 90 187 55 399 102
64 249 344 32 5 7 23 7 220 469 203 306 121 171 342
233 149 208 196 132 1120 18 33 15 13 3 4 52 21 293
523 25 116 17 20 15 179 42 44 323 56 484 74 757 628

2-й квартал 2010 г.

1477 466 269 186 61 462 1548 40 50 18 13 35 417 60 8
19 32 26 55 56 402 289 175 1111 280 108 141 228 281 1097
764 220 30 227 13 333 193 1695 251 1242 1386 126 393 392 16
22 237 813 1113 82 7 11 8 10 188 1438 511 1164 35 57
42 256 1020 1055 750 69 15 827 13 192 214 79 89 20 439
52 152 98 73 54 30 5 18 7 127 7 14 10 7 8 3

3-й квартал 2010 г.

14 12 762 1002 1061 344 8 25 45 11 195 121 608 14 5
3 6 42 4 36 41 1108 12 42 4 95 8 7 12 129
50 25 273 50 354 73 82 71 49 71 32 498 87 29 25
34 54 7 13 8 447 11 5 22 1257 1542 20 286 183 259
3 105 72 57 42 501 37 98 10 10 400 150 25 342 362
128 42 9 19 500 156 34 14 44 408 564 250 176 231 39 97 48

4-й квартал 2010 г.

33 58 139 185 270 155 36 368 128 111 15 18 5 134 3
623 104 56 41 627 326 139 627 59 50 823 185 1623 236 286
248 330 292 1498 46 15 40 106 125 756 72 194 8 122 7
41 548 264 105 16 574 128 16 555 105 320 654 801 948 141
95 288 139 213 13 50 298 44 7 22 77 101 12 12 3
17 570 1017 13 50 67 75 228 128 115 139 223 203 112 128 17 828

1-й квартал 2011 г.

118 68 633 579 321 64 491 47 9 69 185 88 732 236 286
48 91 200 110 101 690 1659 1094 1231 33 1440 189 1455 1053 139
28 735 356 270 492 483 53 93 112 42 587 7 46 94 76
13 50 19 78 255 335 46 154 501 88 221 247 31 392 227
17 258 57 39 61 551 1697 120 627 526 152 79 135 161 16
9 96 338 42 529 86 1080 1092 31 451 420 17 1089 37 1205

2-й квартал 2011 г.

698 70 65 12 21 46 7 10 386 108 362 46 65 308 96
28 139 290 65 34 138 796 264 159 66 53 22 210 25 10
77 29 10 22 23 22 85 544 19 42 231 504 36 112 209
92 23 327 33 1571 1357 596 561 1594 481 1206 83 181 66 361
93 14 47 46 131 53 305 416 1572 697 14 712 23 1141 1556
189 856 1348 54 465 111 367 330 81 253 806 1291 34 777 422 121

3-й квартал 2011 г.

207 13 288 9 4 1648 1833 1133 1218 373 303 35 136 111 488
147 30 55 25 836 122 83 400 228 104 129 60 906 850 127
39 30 127 6 640 1420 1592 1227 1152 1785 1818 31 1547 1616 1644
308 671 1139 31 59 35 216 87 33 75 22 9 19 16 89
453 675 1294 532 1436 561 90 393 30 351 382 843 1479 60 386
61 79 558 292 25 19 218 9 32 93 15 7 14 104 6 9 973

Продолжение таблицы А.1

4-й квартал 2011 г.

1028 1675 250 645 1644 1497 267 1311 1449 330 17 12 3 10 25
1008 1795 1687 1824 1868 1423 1016 6 8 76 1065 3 14 1047 618
209 14 902 201 82 51 111 92 157 60 1586 650 75 51 36
1051 347 644 154 88 37 12 137 25 7 10 3 1256 1505 1859
67 34 4 206 10 4 7 1480 1501 1092 392 548 836 94 193
530 114 569 54 68 403 15 102 13 47 21 39 6 5 7 12 21

1-й квартал 2012 г.

5 19 4 1177 69 549 89 838 82 13 9 3 9 43 13
18 648 100 16 35 23 99 45 30 25 16 664 532 28 44
40 85 109 535 290 551 1300 255 154 17 121 42 29 13 20
89 757 229 326 139 110 502 60 914 146 938 380 834 522 357
71 202 242 449 127 116 442 60 59 368 167 426 273 76 281
631 95 38 25 11 22 33 423 55 168 562 111 1262 349 32 18

2-й квартал 2012 г.

130 138 482 80 155 71 326 41 565 60 18 19 43 40 16
6 89 259 60 579 1212 1205 1439 1588 1331 341 4 1770 395 13
28 164 349 28 46 41 9 49 38 12 6 15 10 186 63
13 12 5 7 31 47 14 1583 1418 506 184 201 83 95 131
742 184 78 64 12 7 1277 16 19 53 56 73 9 452 440
18 8 25 33 82 314 1154 627 1118 243 184 45 18 346 40 392

3-й квартал 2012 г.

336 9 8 11 235 12 72 52 29 34 170 64 17 995 729
386 167 167 563 1236 692 151 273 62 57 261 362 299 12 27
36 37 278 602 352 36 80 181 37 144 18 34 33 39 44
14 97 25 930 274 50 64 14 73 1153 243 839 686 1455 560
335 904 32 122 205 216 517 833 518 46 26 75 774 491 182
15 30 50 10 39 23 108 44 160 485 171 53 49 287 29 249 45

4-й квартал 2012 г.

4 92 818 31 1665 262 124 54 274 348 1415 68 38 30 43
78 791 406 32 7 93 116 172 274 36 313 1207 28 291 17
7 48 98 292 69 40 601 295 311 6 282 43 130 165 119
117 47 39 53 85 100 317 133 34 42 52 10 5 18 5
3 3 5 6 22 93 77 58 25 30 19 12 70 32 336
86 20 35 12 126 565 691 13 111 72 95 84 15 65 28 12 430

1-й квартал 2013 г.

2 609 613 12 5 3 10 39 14 29 33 33 101 7 398
414 60 836 78 192 38 37 57 162 50 82 8 19 7 46
7 220 267 305 185 42 78 44 1568 436 237 1162 258 1188 161
78 205 1536 253 317 126 387 102 29 128 266 118 24 279 339
52 69 40 37 22 31 9 26 63 67 98 1668 9 16 10
222 133 63 147 78 984 120 75 284 347 103 68 17 11 9

2-й квартал 2013 г.

255 29 1092 630 158 12 136 236 14 296 34 19 93 300 811
84 938 29 80 88 7 19 54 31 104 729 366 50 7 5
18 43 10 39 43 93 70 248 40 33 27 62 15 13 15
10 183 51 65 27 45 615 13 28 104 13 161 496 347 624
495 47 479 690 88 73 21 67 153 144 18 835 1398 255 1507
256 209 219 64 311 568 138 1468 1488 126 35 11 5 11 5 355

Продолжение таблицы А.1

3-й квартал 2013 г.

57 2 2 82 180 134 80 109 42 6 378 308 431 46 63
20 9 4 76 33 17 977 14 56 892 172 527 394 90 76
108 466 233 24 12 11 14 220 108 181 380 245 911 307 282
333 86 16 112 66 568 699 103 65 55 117 111 1086 590 13
266 38 19 272 102 51 572 498 790 349 62 243 35 13 5
45 60 52 44 9 14 3 10 51 13 239 127 142 74 219 47 21

4-й квартал 2013 г.

31 347 374 38 48 663 628 59 8 1037 1137 316 388 160 374
30 184 91 16 9 5 13 3 16 15 27 10 25 559 59
440 120 21 192 53 283 601 280 325 41 464 85 298 57 31
383 353 35 432 33 34 34 159 472 86 46 296 162 18 106
262 567 360 89 627 875 198 83 167 455 280 203 83 1085 155
35 79 448 833 901 36 79 137 40 19 4 4 6 8 32 1281 25

1-й квартал 2014 г.

9 131 1188 89 618 120 128 1586 325 342 88 13 620 55 18
668 51 423 138 121 609 16 597 48 45 9 13 11 16 17
70 29 11 10 11 32 23 48 1197 554 620 762 1263 860 116
41 6 11 6 7 7 14 46 41 215 628 140 913 1678 1047
1413 76 234 35 38 54 24 83 813 107 27 137 46 353 168
7 10 9 22 266 70 44 16 983 707 947 124 746 316 957

2-й квартал 2014 г.

243 548 90 23 343 125 76 1262 136 211 17 16 31 6 4
12 4 21 73 231 28 150 177 109 11 1320 1117 199 1164 118
336 27 89 165 1593 921 820 318 185 248 1444 812 208 933 637
4 4 6 5 63 7 122 11 271 338 452 337 146 208 30
11 13 24 45 182 170 96 26 25 60 33 72 500 26 54
91 326 112 91 63 1292 157 675 237 96 357 82 139 1106 746 816

3-й квартал 2014 г.

1340 11 310 1004 32 302 510 6 168 285 83 65 4 9 73
6 6 32 657 141 107 329 487 209 421 444 77 442 122 473
69 16 13 19 10 12 34 6 12 5 20 173 275 316 576
643 121 27 46 40 11 6 10 21 50 35 10 6 9 23
43 1423 301 1674 551 357 398 55 147 294 79 26 257 517 64
1345 123 78 192 71 27 74 82 168 220 286 22 124 34 105 51 391

4-й квартал 2014 г.

171 156 61 156 263 80 395 775 1260 1300 11 24 206 271 247
645 558 145 895 103 90 46 461 540 15 35 1502 720 1783 67
320 568 446 18 1484 152 166 35 1652 41 563 2 13 4 4
5 5 4 3 4 3 15 3 3 3 3 5 159 487 355
105 193 1147 1070 16 249 102 327 143 5 2 3 2 3 5
2 7 5 4 2 2 2 2 2 4 562 749 124 111 457 6 5

1-й квартал 2015 г.

3 149 266 13 44 3 4 12 24 12 578 199 18 84 21
91 179 29 42 164 35 267 378 396 672 109 28 46 665 881
1067 562 255 853 1672 1630 193 892 460 683 5 22 11 15 16
19 21 19 21 87 16 95 20 32 12 73 637 225 144 320
419 1063 45 134 534 684 1038 16 19 51 87 9 5 29 12
34 292 26 20 227 191 11 15 170 32 373 632 513 74 51

Продолжение таблицы А.1

2-й квартал 2015 г.

142 180 58 123 225 71 521 846 1482 1428 12 22 206 289 248
543 664 250 852 92 74 43 453 469 22 35 1552 1292 1774 50
328 624 416 20 1443 152 176 38 1622 38 577 3 2 3 2
2 3 3 3 3 3 16 3 2 3 3 9 152 490 356
116 203 1117 1039 15 264 99 249 206 5 2 6 2 4 5
3 4 3 5 2 2 2 5 5 6 523 845 246 109 411 7

3-й квартал 2015 г.

263 371 35 832 596 181 83 6 12 19 22 437 70 27 8
92 190 8 90 41 21 32 601 729 362 128 15 108 916 434
719 1419 1627 1009 710 82 558 234 354 201 15 115 981 155 1346
100 462 263 213 1163 249 101 5 11 4 4 11 9 23 86
14 100 3 14 10 8 8 7 5 4 174 547 128 1105 478
170 869 114 19 210 34 46 101 46 344 18 88 8 61 69 59 154

4-й квартал 2015 г.

114 10 33 363 75 168 371 41 32 8 6 5 6 14 7
30 345 58 106 942 332 88 308 600 204 69 97 595 125 512
211 982 56 211 439 1369 827 266 369 87 294 160 432 33 17
31 36 91 563 9 908 77 43 383 7 8 5 0 65 246
266 286 1552 1485 1551 1764 1691 205 115 209 566 166 15 44 46
57 55 12 19 54 289 167 87 127 65 313 104 21 21 1319 7 6

1-й квартал 2016 г.

5 3 3 938 182 191 11 37 17 14 4 9 5 5 4
7 31 39 5 19 31 251 151 212 1410 417 537 730 51 29
98 52 33 151 32 271 110 844 26 104 61 108 141 4 328
542 333 1235 212 17 111 15 12 8 88 11 32 150 269 58
660 1022 1324 350 63 157 15 65 108 9 15 15 22 7 899
1237 201 19 45 101 94 15 11 1247 548 10 112 365 305 156 178

2-й квартал 2016 г.

17 21 27 23 84 170 221 406 98 154 57 380 137 66 103
57 23 8 53 851 67 27 28 12 15 32 96 62 298 13
30 615 1067 250 107 364 392 479 19 32 29 630 114 35 63
268 290 1296 487 101 58 96 30 30 700 9 299 208 15 200
189 108 12 179 72 70 230 144 9 77 406 277 64 26 39
265 121 41 147 132 320 67 23 223 667 93 10 27 13 4 25

3-й квартал 2016 г.

119 63 34 215 18 21 4 256 7 13 92 6 48 24 6
18 9 136 187 28 106 30 13 29 359 33 102 112 492 37
156 82 142 1578 1017 179 160 18 10 140 10 1015 672 3 126
143 692 103 667 102 183 145 284 29 578 546 549 57 200 112
347 140 192 168 72 120 7 567 25 103 1325 580 1032 300 440
27 660 13 9 258 9 1083 13 199 250 106 31 47 593 527 175 797

4-й квартал 2016 г.

32 45 157 176 540 58 31 122 489 290 20 120 622 1358 48
36 490 273 23 212 103 59 249 305 361 142 399 371 155 57
54 616 367 256 391 47 196 1487 286 976 33 8 121 17 11
7 17 9 32 67 13 29 123 99 102 257 40 195 105 512
123 192 927 89 24 116 208 510 102 216 944 48 1403 3 67
128 10 1268 434 58 1467 1514 216 381 42 125 481 415 800 1666 1072 216

Продолжение таблицы А.1

1-й квартал 2017 г.

301 130 6 14 607 446 130 25 388 253 174 14 152 11 71
36 509 130 77 356 239 1172 458 632 607 562 472 91 19 29
551 82 337 1082 571 399 337 154 178 1293 127 255 193 312 285
143 4 9 9 7 6 80 17 333 119 27 535 678 15 297
124 560 192 572 591 1243 1209 1048 134 186 354 98 153 846 144
35 137 851 224 728 211 306 48 17 42 59 241 261 406 34

2-й квартал 2017 г.

758 628 109 44 129 120 90 94 157 1289 453 1143 582 136 12
26 110 34 1612 1178 89 31 11 14 10 116 74 59 147 38
134 17 97 46 161 184 41 1079 1186 55 1306 249 494 91 438
963 121 11 56 307 130 69 68 64 49 35 18 523 33 14
9 18 21 13 31 97 467 6 604 56 822 352 674 197 389
375 344 43 405 1490 73 271 87 119 17 72 147 37 220 61 254

3-й квартал 2017 г.

970 533 13 315 729 504 73 215 38 874 867 54 25 44 329
56 72 340 744 156 28 24 24 129 350 101 393 632 617 327
12 296 579 50 241 196 112 809 412 475 95 59 22 35 151
90 243 162 401 97 30 98 762 53 169 227 127 181 395 269
876 248 640 448 62 45 54 519 100 182 26 51 123 113 70
37 73 37 120 50 1710 63 123 25 144 191 270 20 177 92 19 20

4-й квартал 2017 г.

19 151 96 214 204 449 729 810 261 309 22 38 421 317 114
32 82 667 200 656 1644 953 425 17 1349 170 48 32 21 251
943 150 995 35 68 293 463 171 210 577 88 12 150 207 105
73 626 247 126 883 182 299 1426 752 452 21 25 148 25 230
16 395 19 43 71 7 9 78 35 8 12 35 208 235 922
1744 1732 1812 575 1341 683 54 299 138 163 212 27 55 50 2 10 339

1-й квартал 2018 г.

170 597 1650 556 66 124 92 10 19 37 63 7 24 19 35
9 25 56 13 39 7 368 15 69 240 7 72 145 71 82
381 79 677 335 20 69 13 190 378 11 650 181 52 7 152
203 40 377 398 557 86 6 72 185 15 19 33 237 928 148
1219 122 637 547 24 136 389 434 131 34 8 20 43 16 34
87 218 1017 949 756 107 16 16 6 16 12 5 13 5 5

2-й квартал 2018 г.

8 3 6 17 1119 1177 56 453 1036 249 429 407 945 1110 491
231 92 183 54 29 146 193 136 402 17 3 3 3 5 219
213 203 33 329 712 293 48 314 96 324 221 18 20 409 285
17 36 19 10 70 434 42 472 156 711 1163 1458 654 1147 45
267 498 19 7 202 294 824 222 11 8 7 11 16 58 18
8 47 11 320 52 29 193 39 136 20 221 156 17 9 14 18

3-й квартал 2018 г.

11 27 17 60 13 81 83 158 259 176 9 13 40 97 973
16 4 149 790 230 347 563 108 109 542 359 395 51 215 694
89 403 6 113 109 295 680 1570 253 161 728 51 319 1426 966
177 275 102 54 39 41 47 17 77 45 10 53 203 111 288
46 24 410 536 743 579 860 14 167 623 350 137 59 90 713
227 615 125 37 459 202 645 145 131 99 499 265 58 533 113 49 308

Продолжение таблицы А.1

4-й квартал 2018 г.

18	34	97	141	291	1048	1228	596	129	73	87	6	5	5	21
125	27	366	298	252	784	1276	875	10	8	43	56	96	32	19
218	201	772	253	34	24	52	382	165	207	182	167	220	32	71
312	57	5	99	98	681	120	334	70	119	28	60	73	1357	152
575	135	1469	81	583	44	27	263	13	9	4	3	3	9	5
4	4	3	7	770	84	1704	68	1230	127	1170	5	3	8	8	7	6

1-й квартал 2019 г.

417	9	18	7	7	3	8	383	366	56	114	41	10	5	21
15	7	68	12	17	73	9	6	4	362	7	5	603	189	277
67	132	23	8	22	17	61	609	967	121	5	18	85	108	37
155	124	94	110	580	364	154	28	63	28	251	294	45	38	895
109	684	79	42	1651	746	561	177	56	341	365	69	59	6	56
52	206	42	128	10	176	15	24	74	1196	371	98	654	72	23

2-й квартал 2019 г.

87	16	9	45	11	167	198	34	61	21	162	284	106	127	179
170	61	12	31	6	10	5	117	25	182	929	1469	270	100	170
102	160	316	270	100	65	394	701	91	1470	481	134	115	452	121
14	205	93	83	27	263	323	54	386	81	126	16	208	27	575
52	237	17	4	1227	1165	1081	442	13	17	13	11	13	5	12
14	7	24	18	272	46	137	65	66	26	405	280	93	693	369	65

3-й квартал 2019 г.

22	79	242	228	152	63	384	118	482	608	145	53	558	1269	830	
750	385	533	1098	305	192	266	1031	514	70	591	692	952	328	89	
1003	969	1158	11	37	21	99	52	46	39	22	9	46	30	38	
145	52	17	29	54	38	92	45	177	382	75	1117	550	33	191	
120	196	23	10	4	7	10	31	23	43	5	28	6	9	27	
34	56	29	31	44	128	23	13	490	517	162	495	135	32	134	784	64